

PROBABILITY DISTRIBUTIONS SUPPORTED ON SETS GENERATED BY INFINITE AFFINE TRANSFORMATIONS AND OPTIMAL QUANTIZATION

DOĞAN ÇÖMEZ AND MRINAL KANTI ROYCHOWDHURY

ABSTRACT. Quantization of a probability distribution refers to the idea of estimating a given probability by a discrete probability supported by a finite set. In this paper, a probability distribution is considered which is generated by an infinite system of affine transformations S_{ij} on \mathbb{R}^2 associated with probabilities p_{ij} such that $p_{ij} > 0$ for all $i, j \in \mathbb{N}$ and $\sum_{i,j=1}^{\infty} p_{ij} = 1$. For such a probability measure P , the optimal sets of n -means and the n th quantization error are calculated for every natural number n . In addition, it is shown that the distribution of such a probability measure is same as that of the direct product of the Cantor distribution.

1. INTRODUCTION

Quantization is a destructive process which has been extensively studied in information theory (see [4, 9]). Its purpose is to reduce the cardinality of the representation space, in particular when the input data is real-valued. Formally, a quantizer is a function q mapping d -dimensional vectors in the domain $\Omega \subset \mathbb{R}^d$ into a finite set of vectors $\alpha \subset \mathbb{R}^d$. Each vector $a \in \alpha$ is called a *code vector* or a *codeword*, and the set α of all the codewords is called a *codebook*. The Voronoi region generated by $a \in \alpha$, denoted by $M(a|\alpha)$, is defined to be the set of all points in \mathbb{R}^d which are closer to $a \in \alpha$ than to all other points in α , and the set $\{M(a|\alpha) : a \in \alpha\}$ is called the *Voronoi diagram* or *Voronoi tessellation* of \mathbb{R}^d . A special quantization scheme is given by the Voronoi tessellation which associates with each codeword $a \in \alpha$ its Voronoi region $M(a|\alpha)$. For a given probability distribution P on \mathbb{R}^d we define the centroids or mass center, of the regions $M(a|\alpha)$ for $a \in \alpha$, by

$$a^* = \frac{1}{P(M(a|\alpha))} \int_{M(a|\alpha)} x dP = \frac{\int_{M(a|\alpha)} x dP}{\int_{M(a|\alpha)} dP}.$$

A Voronoi tessellation is called a *centroidal Voronoi tessellation* (CVT) if $a^* = a$, i.e., if the generators are also the centroids of their own Voronoi regions. Let $\|\cdot\|$ denote the Euclidean norm on \mathbb{R}^d for any $d \geq 1$. Then for the finite set α , the error $\int \min_{a \in \alpha} \|x - a\|^2 dP(x)$ is often referred to as the *variance*, *cost*, or *distortion error* for α with respect to the probability measure P , and is denoted by $V(\alpha) := V(P; \alpha)$. The value $\inf\{V(P; \alpha) : \alpha \subset \mathbb{R}^d, \text{card}(\alpha) \leq n\}$ is called the *n th quantization error* for the probability measure P , and is denoted by $V_n := V_n(P)$. Such a set α for which the infimum occurs and contains no more than n points is called an *optimal set of n -means*. It is known that for a continuous probability measure an optimal set of n -means always has exactly n -elements (see [7]). The elements of an optimal set are called *optimal quantizers* or *optimal points*. Of course, this makes sense only if the mean-squared error or the expected squared Euclidean distance $\int \|x\|^2 dP(x)$ is finite (see [1, 5, 6, 7]).

2010 *Mathematics Subject Classification.* 60Exx, 28A80, 94A34.

Key words and phrases. Affine transformations, affine set, affine measure, optimal quantizers, quantization error.

The research of the second author was supported by U.S. National Security Agency (NSA) Grant H98230-14-1-0320.

For a Borel probability measure P on \mathbb{R}^d , an optimal set of n -means forms a CVT with n -means (n -generators) of \mathbb{R}^d ; however, the converse is not true in general (see [3, 16]). A Borel measurable partition $\{A_a : a \in \alpha\}$, where α is an index set, of \mathbb{R}^d is called a Voronoi partition of \mathbb{R}^d if $A_a \subset M(a|\alpha)$ for every $a \in \alpha$. Let us now state the following proposition (see [4, 7]):

Proposition 1.1. Let α be an optimal set of n -means and $a \in \alpha$. Then,

(i) $P(M(a|\alpha)) > 0$, (ii) $P(\partial M(a|\alpha)) = 0$, (iii) $a = E(X : X \in M(a|\alpha))$, and (iv) P -almost surely the set $\{M(a|\alpha) : a \in \alpha\}$ forms a Voronoi partition of \mathbb{R}^d .

A transformation $S : X \rightarrow X$ on a metric space (X, d) is called a *contractive* or a *contraction mapping* if there is a constant $0 < c < 1$ such that $d(S(x), S(y)) \leq cd(x, y)$ for all $x, y \in X$. On the other hand, S is called a *similarity mapping* or a *similitude* if there exists a constant $s > 0$ such that $d(S(x), S(y)) = sd(x, y)$ for all $x, y \in X$. Here s is called the *similarity ratio* or the *similarity constant* of the similarity mapping S . Let $P := \frac{1}{2}P \circ S_1^{-1} + \frac{1}{2}P \circ S_2^{-1}$ where $S_1(x) = \frac{1}{3}x$ and $S_2(x) = \frac{1}{3}x + \frac{2}{3}$ for all $x \in \mathbb{R}$. Then, P is a probability distribution on \mathbb{R} with support the Cantor set generated by the similitudes S_1 and S_2 . For such a probability measure Graf and Luschgy determined the optimal sets of n -means and the n th quantization error (see [8]). L. Roychowdhury extended the above result of Graf-Luschgy to probability distributions supported by nonhomogeneous Cantor sets (see [13]). Let us now consider a Sierpiński carpet which is generated by the four contractive similarity mappings S_1, S_2, S_3 and S_4 on \mathbb{R}^2 such that $S_1(x_1, x_2) = \frac{1}{3}(x_1, x_2)$, $S_2(x_1, x_2) = \frac{1}{3}(x_1, x_2) + (\frac{2}{3}, 0)$, $S_3(x_1, x_2) = \frac{1}{3}(x_1, x_2) + (0, \frac{2}{3})$, and $S_4(x_1, x_2) = \frac{1}{3}(x_1, x_2) + (\frac{2}{3}, \frac{2}{3})$ for all $(x_1, x_2) \in \mathbb{R}^2$. If P is a Borel probability measure on \mathbb{R}^2 such that $P = \frac{1}{4}P \circ S_1^{-1} + \frac{1}{4}P \circ S_2^{-1} + \frac{1}{4}P \circ S_3^{-1} + \frac{1}{4}P \circ S_4^{-1}$, then P has support the Sierpiński carpet. For this probability measure, Çömez and Roychowdhury determined the optimal sets of n -means and the n th quantization error (see [2]). Let us now consider a probability measure P on \mathbb{R} which is generated by an infinite collection of similitudes $\{S_j\}_{j=1}^{\infty}$ such that $S_j(x) = \frac{1}{3^j}x + 1 - \frac{1}{3^{j-1}}$ for all $x \in \mathbb{R}$ and P is given by $P = \sum_{j=1}^{\infty} \frac{1}{2^j}P \circ S_j^{-1}$. For this probability measure, Roychowdhury determined the optimal sets of n -means and the n th quantization error (see [14]), and this result is an infinite extension of the result of Graf and Luschgy in [8]. For an infinite extension of the result of L. Roychowdhury see [15]. In this paper, we made an infinite extension of the result of Çömez and Roychowdhury in [2].

Let $\{S_{(i,j)} : i, j \in \mathbb{N}\}$ be a collection of infinite affine transformations on \mathbb{R}^2 such that $S_{(i,j)}(x_1, x_2) = (\frac{1}{3^i}x_1 + 1 - \frac{1}{3^{i-1}}, \frac{1}{3^j}x_2 + 1 - \frac{1}{3^{j-1}})$. Clearly these affine transformations are all contractive but not similarity mappings. Let us now associate the mappings $S_{(i,j)}$ with the probabilities $p_{(i,j)}$ such that $p_{(i,j)} = \frac{1}{2^{i+j}}$ for all $i, j \in \mathbb{N}$, where $\mathbb{N} := \{1, 2, 3, \dots\}$ is the set of all natural numbers. Then, there exists a unique Borel probability measure P on \mathbb{R}^2 (see [10], [12], [11], etc.) such that

$$P = \sum_{i,j=1}^{\infty} p_{(i,j)} P \circ S_{(i,j)}^{-1}.$$

The support of such a probability measure lies in the unit square $[0, 1] \times [0, 1]$. We call such a measure an *affine measure* on \mathbb{R}^2 or more specifically an *infinitely generated affine measure* on \mathbb{R}^2 . This paper deals with this measure P . The arrangement of the paper is as follows: In Section 2, we discuss the basic definitions and lemmas about the optimal sets of n -means and the n th quantization error. In Section 3, we determine the optimal sets of n -means for $n = 2$ and $n = 3$. In Section 4, first we define a mapping F which helps us to convert the infinitely generated affine measure P to a finitely generated product measure $P_c \times P_c$ where each P_c is the Cantor distribution given by $P_c = \frac{1}{2}P \circ U_1^{-1} + \frac{1}{2}P_c \circ U_2^{-1}$ where $U_1(x) = \frac{1}{3}x$ and $U_2(x) = \frac{1}{3}x + \frac{2}{3}$ for all $x \in \mathbb{R}$. Section 5 mainly contains the main result of the paper:

Propositions 5.6, 5.10 and 5.14 give the closed formulas to determine the optimal sets of n -means and the corresponding quantization error for all $n \geq 4$. We also give some examples and figures to illustrate the constructions further.

2. PRELIMINARIES

In this section we give the basic definitions and the preliminary lemmas that will be instrumental in determining the optimal sets of n -means. Let P be the affine measure on \mathbb{R}^2 as defined before generated by the contractive affine transformations given by $S_{(i,j)}(x_1, x_2) = (\frac{1}{3^i}x_1 + 1 - \frac{1}{3^{i-1}}, \frac{1}{3^j}x_2 + 1 - \frac{1}{3^{j-1}})$ for all $(x_1, x_2) \in \mathbb{R}^2$ and $i, j \in \mathbb{N}$. Let us consider the alphabet $I := \mathbb{N}^2 = \{(i, j) : i, j \in \mathbb{N}\}$. By a *string* or a *word* ω over the alphabet I , it is meant a finite sequence $\omega := \omega_1\omega_2 \cdots \omega_k$ of symbols from the alphabet, where $k \geq 1$, and k is called the length of the word ω . A word of length zero is called the *empty word*, and is denoted by \emptyset . By I^* we denote the set of all words over the alphabet I of some finite length k including the empty word \emptyset . By $|\omega|$, we denote the length of a word $\omega \in I^*$. For any two words $\omega := \omega_1\omega_2 \cdots \omega_k$ and $\tau := \tau_1\tau_2 \cdots \tau_\ell$ in I^* , by $\omega\tau := \omega_1 \cdots \omega_k\tau_1 \cdots \tau_\ell$ we mean the word obtained from the concatenation of ω and τ . For $n \geq 1$ and $\omega = \omega_1\omega_2 \cdots \omega_n \in I^*$ we define $\omega^- := \omega_1\omega_2 \cdots \omega_{n-1}$. Note that ω^- is the empty word if the length of ω is one. Analogously, by \mathbb{N}^* we denote the set of all words over the alphabet \mathbb{N} , and for any $\tau \in \mathbb{N}^*$ similar is the meaning for $|\tau|$, τ^- , etc. Let $\omega \in I^k$, $k \geq 1$, be such that $\omega = (i_1, j_1)(i_2, j_2) \cdots (i_k, j_k)$, then $\omega^{(1)}$ and $\omega^{(2)}$ will denote $\omega^{(1)} := i_1i_2 \cdots i_k$ and $\omega^{(2)} := j_1j_2 \cdots j_k$. Thus, $\omega_{|\omega|}^{(1)} = i_k$ and $\omega_{|\omega|}^{(2)} = j_k$. These lead us to define the following notations: For $\omega \in I^*$, by $\omega(\emptyset, \infty)$ it is meant the set of all words $\omega^-(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)} + j)$ obtained by concatenating the word ω^- with the word $(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)} + j)$ for $j \in \mathbb{N}$, i.e.,

$$\omega(\emptyset, \infty) := \{\omega^-(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)} + j) : j \in \mathbb{N}\}.$$

Similarly, $\omega(\infty, \emptyset)$ and $\omega(\infty, \infty)$ represent the sets

$$\omega(\infty, \emptyset) := \{\omega^-(\omega_{|\omega|}^{(1)} + i, \omega_{|\omega|}^{(2)}) : i \in \mathbb{N}\} \text{ and } \omega(\infty, \infty) := \{\omega^-(\omega_{|\omega|}^{(1)} + i, \omega_{|\omega|}^{(2)} + j) : i, j \in \mathbb{N}\}$$

respectively. Analogously, for any $\tau \in \mathbb{N}^*$, by (τ, ∞) it is meant the set $(\tau, \infty) := \{\tau + i : i \in \mathbb{N}\}$, and (τ, \emptyset) represents the set $(\tau, \emptyset) := \{\tau\}$. Thus, if $\omega = (i_1, j_1)(i_2, j_2) \cdots (i_k, j_k)(\infty, \emptyset)$, then we write $\omega^{(1)} := (i_1i_2 \cdots i_k, \infty)$ and $\omega^{(2)} := j_1j_2 \cdots j_k$; if $\omega = (i_1, j_1)(i_2, j_2) \cdots (i_k, j_k)(\emptyset, \infty)$, then we write $\omega^{(1)} := i_1i_2 \cdots i_k$ and $\omega^{(2)} := (j_1j_2 \cdots j_k, \infty)$; and if $\omega = (i_1, j_1)(i_2, j_2) \cdots (i_k, j_k)(\infty, \infty)$, then we write $\omega^{(1)} := (i_1i_2 \cdots i_k, \infty)$ and $\omega^{(2)} := (j_1j_2 \cdots j_k, \infty)$. For $\omega = \omega_1\omega_2 \cdots \omega_k \in I^k$, $k \geq 1$, let us write

$$S_\omega := S_{\omega_1} \circ \cdots \circ S_{\omega_k}, \quad p_\omega := p_{\omega_1}p_{\omega_2} \cdots p_{\omega_k} \text{ and } J_\omega := S_\omega([0, 1] \times [0, 1]).$$

If ω is the empty word \emptyset , by S_ω we mean the identity mapping on \mathbb{R}^2 and write $J := J_\emptyset = S_\emptyset([0, 1] \times [0, 1])$. Then, the probability measure P has support the closure of the limit set S , where $S = \bigcap_{k \in \mathbb{N}} \bigcup_{\omega \in I^k} J_\omega$. The limit set S is called the *affine set* or more specifically *infinitely generated affine set*. For $\omega \in I^*$ and $i, j \in \mathbb{N}$, the rectangles $J_{\omega(i,j)}$, into which J_ω is split up at the $(k+1)$ th level are called the *children* or the *basic rectangles* of J_ω (see Figure 1). For $\omega \in I^*$, we write

$$J_{\omega(\emptyset, \infty)} := \bigcup_{j=1}^{\infty} J_{\omega^-(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)} + j)}, \quad J_{\omega(\infty, \emptyset)} := \bigcup_{i=1}^{\infty} J_{\omega^-(\omega_{|\omega|}^{(1)} + i, \omega_{|\omega|}^{(2)})}, \quad J_{\omega(\infty, \infty)} := \bigcup_{i,j=1}^{\infty} J_{\omega^-(\omega_{|\omega|}^{(1)} + i, \omega_{|\omega|}^{(2)} + j)},$$

$$p_{\omega(\emptyset, \infty)} := P(J_{\omega(\emptyset, \infty)}) = \sum_{j=1}^{\infty} p_{\omega^-(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)} + j)},$$

$$p_{\omega(\infty, \emptyset)} := P(J_{\omega(\infty, \emptyset)}) = \sum_{i=1}^{\infty} p_{\omega^-(\omega_{|\omega|}^{(1)} + i, \omega_{|\omega|}^{(2)})}, \quad p_{\omega(\infty, \infty)} := P(J_{\omega(\infty, \infty)}) = \sum_{i,j=1}^{\infty} p_{\omega^-(\omega_{|\omega|}^{(1)} + i, \omega_{|\omega|}^{(2)} + j)}.$$

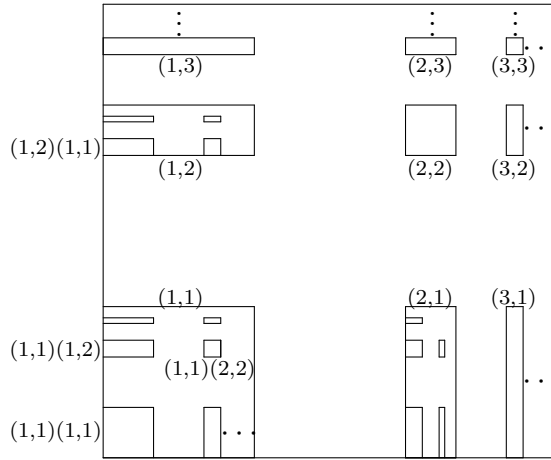


FIGURE 1. Basic rectangles of the infinite affine transformations.

Notice that for any $\omega \in I^*$, $p_\omega(\emptyset, \infty) = p_\omega - \sum_{j=1}^{\infty} \frac{1}{2^{\omega_{| \omega |}^{(1)} + \omega_{| \omega |}^{(2)} + j}} = p_\omega - p_{\omega_{| \omega |}} \sum_{j=1}^{\infty} \frac{1}{2^j} = p_\omega - p_{\omega_{| \omega |}} = p_\omega$; and similarly one can see that $p_\omega(\infty, \emptyset) = p_\omega(\infty, \infty) = p_\omega$.

Lemma 2.1. *Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ be Borel measurable and $k \in \mathbb{N}$. Then,*

$$\int f dP = \sum_{\omega \in I^k} p_\omega \int f \circ S_\omega dP.$$

Proof. We know $P = \sum_{i,j=1}^{\infty} p_{(i,j)} P \circ S_{(i,j)}^{-1}$, and so by induction $P = \sum_{\omega \in I^k} p_\omega P \circ S_\omega^{-1}$, and thus the lemma is yielded. \square

Let $S_{(i,j)}^{(1)}$ and $S_{(i,j)}^{(2)}$ be the horizontal and vertical components of the transformations $S_{(i,j)}$ for all $i, j \in \mathbb{N}$. Then for all $(x_1, x_2) \in \mathbb{R}^2$ we have $S_{(i,j)}^{(1)}(x_1) = \frac{1}{3^i}x_1 + 1 - \frac{1}{3^{i-1}}$ and $S_{(i,j)}^{(2)}(x_2) = \frac{1}{3^j}x_2 + 1 - \frac{1}{3^{j-1}}$. Thus, we see that $S_{(i,j)}^{(1)}$ and $S_{(i,j)}^{(2)}$ are all similarity mappings on \mathbb{R} . Let their similarity ratios be denoted respectively by $s_{(i,j)}^{(1)}$ and $s_{(i,j)}^{(2)}$. Then, $s_{(i,j)}^{(1)} = \frac{1}{3^i}$ and $s_{(i,j)}^{(2)} = \frac{1}{3^j}$. Similarly, for $\omega = (i_1, j_1)(i_2, j_2) \cdots (i_k, j_k) \in I^k$, $k \geq 1$, let $S_\omega^{(1)}$ and $S_\omega^{(2)}$ represent the horizontal and vertical components of the transformation S_ω on \mathbb{R}^2 . Then, $S_\omega^{(1)}$ and $S_\omega^{(2)}$ are similarity mappings on \mathbb{R} with similarity ratios $s_\omega^{(1)}$ and $s_\omega^{(2)}$ respectively, such that $S_\omega^{(1)} = S_{(i_1, j_1)}^{(1)} \circ \cdots \circ S_{(i_k, j_k)}^{(1)}$ and $S_\omega^{(2)} = S_{(i_1, j_1)}^{(2)} \circ \cdots \circ S_{(i_k, j_k)}^{(2)}$. Thus, one can see that

$$s_\omega^{(1)} = s_{(i_1, j_1)}^{(1)} s_{(i_2, j_2)}^{(1)} \cdots s_{(i_k, j_k)}^{(1)} = \left(\frac{1}{3}\right)^{i_1 + i_2 + \cdots + i_k} \quad \text{and}$$

$$s_\omega^{(2)} = s_{(i_1, j_1)}^{(2)} s_{(i_2, j_2)}^{(2)} \cdots s_{(i_k, j_k)}^{(2)} = \left(\frac{1}{3}\right)^{j_1 + j_2 + \cdots + j_k}.$$

Moreover, we have $P(J_\omega) = p_\omega = p_{(i_1, j_1)} p_{(i_2, j_2)} \cdots p_{(i_k, j_k)} = \frac{1}{2^{i_1 + i_2 + \cdots + i_k + j_1 + j_2 + \cdots + j_k}}$. Let $X = (X_1, X_2)$ be a bivariate random variable with distribution P . Let P_1, P_2 be the marginal distributions of P , i.e., $P_1(A) = P(A \times \mathbb{R}) = P \circ \pi_1^{-1}(A)$ for all $A \in \mathfrak{B}$, and $P_2(B) = P(\mathbb{R} \times B) = P \circ \pi_2^{-1}(B)$ for all $B \in \mathfrak{B}$, where π_1, π_2 are projections given by $\pi_1(x_1, x_2) = x_1$ and $\pi_2(x_1, x_2) = x_2$ for all $(x_1, x_2) \in \mathbb{R}^2$. Here \mathfrak{B} is the Borel σ -algebra on \mathbb{R} . Then X_1 has distribution P_1 and X_2 has distribution P_2 . Let $S_{(i,j)}^{-(1)}$ and $S_{(i,j)}^{-(2)}$ denote respectively the inverse images of the horizontal and vertical components of the transformations $S_{(i,j)}$ for all $i, j \in \mathbb{N}$.

Then, the following lemma is known (see [10], [12], [11], etc.).

Lemma 2.2. *Let P_1 and P_2 be the marginal distributions of the probability measure P . Then,*

$$P_1 = \sum_{i=1}^{\infty} \frac{1}{2^i} P_1 \circ S_{(i,j)}^{-(1)} \text{ and } P_2 = \sum_{j=1}^{\infty} \frac{1}{2^j} P_2 \circ S_{(i,j)}^{-(2)}.$$

Remark 2.3. Since $S_{(i,j)}^{(1)}$ and $S_{(i,j)}^{(2)}$ are similarity mappings, from Lemma 2.2, one can see that both the marginal distributions P_1 and P_2 are self-similar measures on \mathbb{R} generated by an infinite collection of similarity mappings associated with the probability vector $(\frac{1}{2}, \frac{1}{2^2}, \dots)$. Recall that for such a probability measure Roychowdhury determined the optimal sets of n -means and the n th quantization error for every natural number n (see [14]). In the sequel, alternatively we will write T_i for $S_{(i,j)}^{(1)}$, and T_j for $S_{(i,j)}^{(2)}$, where T_k for all $k \geq 1$ form an infinite collection of similarity mappings on \mathbb{R} such that $T_k(x) = \frac{1}{3^k}x + 1 - \frac{1}{3^{k-1}}$ for all $x \in \mathbb{R}$. Thus, if $\omega = (i_1, j_1)(i_2, j_2) \cdots (i_n, j_n)$, then $S_{\omega}^{(1)} = T_{i_1} \circ \cdots \circ T_{i_n} = T_{i_1 i_2 \cdots i_n}$ and $S_{\omega}^{(2)} = T_{j_1} \circ \cdots \circ T_{j_n} = T_{j_1 j_2 \cdots j_n}$ for all $n \geq 1$. Again, T_{\emptyset} is the identity mapping on \mathbb{R} .

Lemma 2.4. *Let $E(X)$ and $V(X)$ denote the expectation and the variance of the random variable X . Then,*

$$E(X) = (E(X_1), E(X_2)) = (\frac{1}{2}, \frac{1}{2}) \text{ and } V := V(X) = E\|X - (\frac{1}{2}, \frac{1}{2})\|^2 = \frac{1}{4}.$$

Proof. By Lemma 2.2, one can see that if P_1 and P_2 are the marginal distributions of the probability measure P , then $P_1 = P_2 = \mu$, where μ is a unique Borel probability measure on \mathbb{R} such that

$$\mu = \sum_{k=1}^{\infty} \frac{1}{2^k} \mu \circ T_k^{-1},$$

where T_k are the mappings as defined in Remark 2.3 associated with the probability vector $(\frac{1}{2}, \frac{1}{2^2}, \dots)$. Hence, $X_1 = X_2$, and by [13, Lemma 2.2], we have

$$E(X_1) = E(X_2) = \frac{1}{2}, \text{ and } V(X_1) = V(X_2) = \frac{1}{8},$$

which implies that

$$E\|X - (\frac{1}{2}, \frac{1}{2})\|^2 = E(X_1 - \frac{1}{2})^2 + E(X_2 - \frac{1}{2})^2 = V(X_1) + V(X_2) = \frac{1}{4}.$$

Hence, the lemma follows. \square

Remark 2.5. Using the standard rule of probability, for any $(a, b) \in \mathbb{R}^2$, we have $E\|X - (a, b)\|^2 = V + \|(a, b) - (\frac{1}{2}, \frac{1}{2})\|^2$, which yields that the optimal set of one-mean consists of the expected value and the corresponding quantization error is the variance V of the random variable X .

Lemma 2.6. *Let $\omega \in I^*$. Then,*

- (i) $E(X|X \in J_{\omega(\infty, \infty)}) = S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}(\frac{1}{2}, \frac{1}{2}) + (s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}, s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)});$
- (ii) $E(X|X \in J_{\omega(\emptyset, \infty)}) = S_{\omega^-(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)}+1)}(\frac{1}{2}, \frac{1}{2}) + (0, s_{\omega^-(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)}+1)}^{(2)}), \text{ and}$
- (iii) $E(X|X \in J_{\omega(\infty, \emptyset)}) = S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)})}(\frac{1}{2}, \frac{1}{2}) + (s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)})}^{(1)}, 0).$

Proof. Let us first prove (i). Recall that $P(J_{\omega(\infty, \infty)}) = p_{\omega(\infty, \infty)} = p_\omega$ and $p_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)} = p_\omega \frac{1}{2^{i+j}}$. Then, by the definition of conditional expectation, we have

$$\begin{aligned} E(X|X \in J_{\omega(\infty, \infty)}) &= E(X|X \in \bigcup_{i,j=1}^{\infty} J_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}) \\ &= \frac{1}{P(J_{\omega(\infty, \infty)})} \sum_{i,j=1}^{\infty} p_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)} S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}\left(\frac{1}{2}, \frac{1}{2}\right) = \sum_{i,j=1}^{\infty} \frac{1}{2^{i+j}} S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}\left(\frac{1}{2}, \frac{1}{2}\right). \end{aligned}$$

Notice that

$$\begin{aligned} &S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}\left(\frac{1}{2}, \frac{1}{2}\right) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}\left(\frac{1}{2}, \frac{1}{2}\right) \\ &= \left(S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}\right), S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}\left(\frac{1}{2}\right)\right) - \left(S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1}{2}\right), S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}\left(\frac{1}{2}\right)\right) \\ &= \left(S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}\right) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1}{2}\right), S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}\left(\frac{1}{2}\right) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}\left(\frac{1}{2}\right)\right). \end{aligned}$$

Since

$$\begin{aligned} &S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}\right) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1}{2}\right) = s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(S_{\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j}^{(1)}\left(\frac{1}{2}\right) - S_{\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1}^{(1)}\left(\frac{1}{2}\right)\right) \\ &= s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{3^{\omega_{|\omega|}^{(1)}+i}} - \frac{1}{3^{\omega_{|\omega|}^{(1)}+i-1}} - \frac{1}{3^{\omega_{|\omega|}^{(1)}+1}} \frac{1}{2} + \frac{1}{3^{\omega_{|\omega|}^{(1)}+1-1}}\right) = s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2} \frac{1}{3^i} - \frac{1}{3^{i-1}} - \frac{1}{6} + 1\right) \\ &= s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{5}{6} - \frac{5}{2} \frac{1}{3^i}\right), \text{ and similarly} \end{aligned}$$

$$S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}\left(\frac{1}{2}\right) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}\left(\frac{1}{2}\right) = s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}\left(\frac{5}{6} - \frac{5}{2} \frac{1}{3^j}\right), \text{ we have}$$

$$S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}\left(\frac{1}{2}, \frac{1}{2}\right) = S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}\left(\frac{1}{2}, \frac{1}{2}\right) + (s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{5}{6} - \frac{5}{2} \frac{1}{3^i}\right), s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}\left(\frac{5}{6} - \frac{5}{2} \frac{1}{3^j}\right)).$$

Thus, we deduce

$$\begin{aligned} E(X|X \in J_{\omega(\infty, \infty)}) &= S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}\left(\frac{1}{2}, \frac{1}{2}\right) + \sum_{i,j=1}^{\infty} \frac{1}{2^{i+j}} (s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{5}{6} - \frac{5}{2} \frac{1}{3^i}\right), s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}\left(\frac{5}{6} - \frac{5}{2} \frac{1}{3^j}\right)) \\ &= S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}\left(\frac{1}{2}, \frac{1}{2}\right) + (s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)} \frac{1}{3}, s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)} \frac{1}{3}) \\ &= S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}\left(\frac{1}{2}, \frac{1}{2}\right) + (s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}, s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}). \end{aligned}$$

The last equation in the above expression follows from the fact that

$$s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)} = s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)} s_{\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1}^{(1)} = s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)} \frac{1}{3} = s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)} \frac{1}{3},$$

and $s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)} = s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)} \frac{1}{3}$, which is obtained similarly.

Likewise, one can prove (ii) and (iii). Thus, the proof of the lemma is complete. \square

Note 2.7. For words $\beta, \gamma, \dots, \delta$ in I^* , by $a(\beta, \gamma, \dots, \delta)$ we denote the conditional expectation of the random variable X given $J_\beta \cup J_\gamma \cup \dots \cup J_\delta$, i.e.,

$$(1) \quad a(\beta, \gamma, \dots, \delta) = E(X|X \in J_\beta \cup J_\gamma \cup \dots \cup J_\delta) = \frac{1}{P(J_\beta \cup \dots \cup J_\delta)} \int_{J_\beta \cup \dots \cup J_\delta} (x_1, x_2) dP.$$

Then, for $\omega \in I^*$,

$$(2) \quad \begin{cases} a(\omega) = S_\omega(E(X)) = S_\omega(\frac{1}{2}, \frac{1}{2}), & a(\omega(\emptyset, \infty)) = E(X|X \in J_{\omega(\emptyset, \infty)}), \\ a(\omega(\infty, \emptyset)) = E(X|X \in J_{\omega(\infty, \emptyset)}) \text{ and } a(\omega(\infty, \infty)) = E(X|X \in J_{\omega(\infty, \infty)}). \end{cases}$$

Thus, by Lemma 2.6, if $\omega = (1, 1)$, then $a((1, 1)) = (\frac{1}{6}, \frac{1}{6})$, $a((1, 1)(\infty, \emptyset)) = (\frac{5}{6}, \frac{1}{6})$, $a((1, 1)(\emptyset, \infty)) = (\frac{1}{6}, \frac{5}{6})$, and $a((1, 1)(\infty, \infty)) = (\frac{5}{6}, \frac{5}{6})$. In addition,

$$(3) \quad \begin{cases} a((1, 1), (1, 1)(\infty, \emptyset)) = (\frac{1}{2}, \frac{1}{6}), & a((1, 1)(\emptyset, \infty), (1, 1)(\infty, \infty)) = (\frac{1}{2}, \frac{5}{6}), \\ a((1, 1), (1, 1)(\emptyset, \infty)) = (\frac{1}{6}, \frac{1}{2}), & a((1, 1)(\infty, \emptyset), (1, 1)(\infty, \infty)) = (\frac{5}{6}, \frac{1}{2}). \end{cases}$$

Moreover, for $\omega \in I^k$, $k \geq 1$, it is easy to see that

$$(4) \quad \begin{aligned} \int_{J_\omega} \|x - (a, b)\|^2 dP &= p_\omega \int \|(x_1, x_2) - (a, b)\|^2 dP \circ S_\omega^{-1} \\ &= p_\omega \left(s_\omega^{(1)2} V(X_1) + s_\omega^{(2)2} V(X_2) + \|S_\omega(\frac{1}{2}, \frac{1}{2}) - (a, b)\|^2 \right), \end{aligned}$$

where $s_\omega^{(k)2} := (s_\omega^{(k)})^2$ for $k = 1, 2$. The expressions (2) and (4) are useful to obtain the optimal sets and the corresponding quantization errors with respect to the probability distribution P .

Lemma 2.8. *Let P be the affine measure on \mathbb{R}^2 and let $\omega \in I^*$. Then,*

$$\begin{aligned} \int_{J_{\omega(\infty, \infty)}} \|x - a(\omega(\infty, \infty))\|^2 dP &= \int_{J_{\omega(\emptyset, \infty)}} \|x - a(\omega(\emptyset, \infty))\|^2 dP \\ &= \int_{J_{\omega(\infty, \emptyset)}} \|x - a(\omega(\infty, \emptyset))\|^2 dP = \int_{J_\omega} \|x - a(\omega)\|^2 dP = p_\omega (s_\omega^{(1)2} + s_\omega^{(2)2}) \frac{1}{8}. \end{aligned}$$

Proof. Let us first prove $\int_{J_{\omega(\infty, \infty)}} \|x - a(\omega(\infty, \infty))\|^2 dP = p_\omega (s_\omega^{(1)2} + s_\omega^{(2)2}) \frac{1}{8}$. By Lemma 2.6, we have

$$\begin{aligned} (5) \quad \int_{J_{\omega(\infty, \infty)}} \|x - a(\omega(\infty, \infty))\|^2 dP &= \sum_{i,j=1}^{\infty} \int_{J_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}} \|x - a(\omega(\infty, \infty))\|^2 dP \\ &= p_\omega \sum_{i,j=1}^{\infty} \frac{1}{2^{i+j}} \int \|S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}(x_1, x_2) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}(\frac{1}{2}, \frac{1}{2}) \\ &\quad - (s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}, s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)})\|^2 dP. \end{aligned}$$

Note that $S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}(x_1, x_2) = \left(S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}(x_1), S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}(x_2) \right)$ and

$S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}(\frac{1}{2}, \frac{1}{2}) = \left(S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}(\frac{1}{2}), S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}(\frac{1}{2}) \right)$. Moreover, we have

$$\begin{aligned} &\left(S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}(x_1) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}(\frac{1}{2}) - s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)} \right)^2 \\ &= s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)2} \left(S_{\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j}^{(1)}(x_1) - S_{\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1}^{(1)}(\frac{1}{2}) - s_{\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1}^{(1)} \right)^2 \\ &= s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)2} \left(\left(S_{\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j}^{(1)}(x_1) - S_{\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j}^{(1)}(\frac{1}{2}) \right) + \left(S_{\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j}^{(1)}(\frac{1}{2}) - S_{\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1}^{(1)}(\frac{1}{2}) \right) \right. \\ &\quad \left. - s_{\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1}^{(1)} \right) \right)^2. \end{aligned}$$

Now break the above expression using the square formula and note the fact that

$$\begin{aligned} \int \left(S_{(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}(x_1) - S_{(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}\right) \right)^2 dP_1 &= s_{(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)2} V(X_1) = s_{(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)})}^{(1)2} \frac{1}{9^i} \frac{1}{8}, \text{ and} \\ \int \left(S_{(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}(x_1) - S_{(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}\right) \right) dP_1 &= 0, \text{ and after some simplification we have} \\ \left(S_{(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}\left(\frac{1}{2}\right) - S_{(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1}{2}\right) - s_{(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)} \right)^2 &= s_{(\omega_{|\omega|}^{(1)}, \omega_{|\omega|}^{(2)})}^{(1)2} \frac{1}{4} \left(1 - \frac{5}{3^i}\right)^2. \end{aligned}$$

Thus, it follows that

$$\begin{aligned} &\int \left(S_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)}(x_1) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}\left(\frac{1}{2}\right) - s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)} \right)^2 dP_1 \\ &= s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)2} \left(\frac{1}{9^i} \frac{1}{8} + \frac{1}{4} \left(1 - \frac{5}{3^i}\right)^2 \right), \text{ and similarly} \\ &\int \left(S_{\omega^-(\omega_{|\omega|}^{(2)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)}(x_2) - S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}\left(\frac{1}{2}\right) - s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)} \right)^2 dP_2 \\ &= s_{\omega^-(\omega_{|\omega|}^{(2)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)2} \left(\frac{1}{9^j} \frac{1}{8} + \frac{1}{4} \left(1 - \frac{5}{3^j}\right)^2 \right). \end{aligned}$$

Therefore, (5) implies that

$$\begin{aligned} &\int_{J_{\omega(\infty, \infty)}} \|x - a(\omega(\infty, \infty))\|^2 dP \\ &= p_{\omega} \sum_{i,j=1}^{\infty} \frac{1}{2^{i+j}} \left(s_{\omega^-(\omega_{|\omega|}^{(1)}+i, \omega_{|\omega|}^{(2)}+j)}^{(1)2} \left(\frac{1}{9^i} \frac{1}{8} + \frac{1}{4} \left(1 - \frac{5}{3^i}\right)^2 \right) + s_{\omega^-(\omega_{|\omega|}^{(2)}+i, \omega_{|\omega|}^{(2)}+j)}^{(2)2} \left(\frac{1}{9^j} \frac{1}{8} + \frac{1}{4} \left(1 - \frac{5}{3^j}\right)^2 \right) \right) = p_{\omega} (s_{\omega}^{(1)2} + s_{\omega}^{(2)2}) \frac{1}{8}. \end{aligned}$$

Similarly, one can prove the other parts of the lemma. Thus, the proof of the lemma is complete. \square

3. OPTIMAL SETS OF n -MEANS FOR $n = 2, 3$

In this section, we determine the optimal sets of two- and three-means, and their quantization errors.

Lemma 3.1. *Let P be the affine measure on \mathbb{R}^2 , and let $\{(a, p), (b, p)\}$ be a set of two points lying on the line $x_2 = p$ for which the distortion error is smallest. Then, $a = \frac{1}{6}$, $b = \frac{5}{6}$, $p = \frac{1}{2}$ and the distortion error is $\frac{5}{36} = 0.138889$.*

Proof. Let $\beta = \{(a, p), (b, p)\}$. Since the points for which the distortion error is smallest are the centroids of their own Voronoi regions, by the properties of centroids, we have

$$(a, p)P(M((a, p)|\beta)) + (b, p)P(M((b, p)|\beta)) = \left(\frac{1}{2}, \frac{1}{2}\right),$$

which implies $pP(M((a, p)|\beta)) + pP(M((b, p)|\beta)) = \frac{1}{2}$, i.e., $p = \frac{1}{2}$. Thus, the boundary of the Voronoi regions is the line $x_1 = \frac{1}{2}$. Now, using the definition of conditional expectation,

$$\left(a, \frac{1}{2}\right) = E(X : X \in M((a, \frac{1}{2})|\beta)) = E(X : X \in \bigcup_{j=1}^{\infty} J_{(1,j)}) = \frac{1}{\sum_{j=1}^{\infty} p_{(1,j)}} \sum_{j=1}^{\infty} p_{(1,j)} S_{(1,j)}\left(\frac{1}{2}, \frac{1}{2}\right),$$

which implies $(a, \frac{1}{2}) = (\frac{1}{6}, \frac{1}{2})$ yielding $a = \frac{1}{6}$. Similarly, $b = \frac{5}{6}$. Then, the distortion error is

$$\int \min_{c \in \beta} \|x - c\|^2 dP = \int_{\bigcup_{j=1}^{\infty} J_{(1,j)}} \|x - (\frac{1}{6}, \frac{1}{2})\|^2 dP + \int_{\bigcup_{i=2, j=1}^{\infty} J_{(i,j)}} \|x - (\frac{5}{6}, \frac{1}{2})\|^2 dP = \frac{5}{72} + \frac{5}{72} = \frac{5}{36}.$$

This completes the proof the lemma. \square

The following lemma provides us information on where to look for points of an optimal set of two-means.

Lemma 3.2. *Let P be the affine measure on \mathbb{R}^2 . The points in an optimal set of two-means can not lie on an oblique line of the affine set.*

Proof. In the affine set, among all the oblique lines that pass through the point $(\frac{1}{2}, \frac{1}{2})$, the line $x_2 = x_1$ has the maximum symmetry, i.e., with respect to the line $x_2 = x_1$ the affine set is geometrically symmetric. Also, observe that, if the two basic rectangles of similar geometrical shape lie in the opposite sides of the line $x_2 = x_1$, and are equidistant from the line $x_2 = x_1$, then they have the same probability (see Figure 1); hence, they are symmetric with respect to the probability distribution P . Due to this, among all the pairs of two points which have the boundaries of the Voronoi regions oblique lines passing through the point $(\frac{1}{2}, \frac{1}{2})$, the two points which have the boundary of the Voronoi regions the line $x_2 = x_1$ will give the smallest distortion error. Again, we know the two points which give the smallest distortion error are the centroids of their own Voronoi regions. Let (a_1, b_1) and (a_2, b_2) be the centroids of the left half and the right half of the affine set with respect to the line $x_2 = x_1$ respectively. Then from the definition of conditional expectation, we have

$$\begin{aligned} (a_1, b_1) = & 2 \left[\sum_{i=1, j=i+1}^{\infty} \frac{1}{2^{i+j}} S_{(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) + \sum_{k_1=1}^{\infty} \sum_{\substack{i=1 \\ j=i+1}}^{\infty} \frac{1}{2^{2k_1+i+j}} S_{(k_1, k_1)(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) \right. \\ & + \sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} \sum_{\substack{i=1 \\ j=i+1}}^{\infty} \frac{1}{2^{2k_1+2k_2+i+j}} S_{(k_1, k_1)(k_2, k_2)(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) \\ & \left. + \sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} \sum_{k_3=1}^{\infty} \sum_{\substack{i=1 \\ j=i+1}}^{\infty} \frac{1}{2^{2k_1+2k_2+2k_3+i+j}} S_{(k_1, k_1)(k_2, k_2)(k_3, k_3)(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) + \dots \right] = \left(\frac{3}{10}, \frac{7}{10} \right), \end{aligned}$$

and

$$\begin{aligned} (a_2, b_2) = & 2 \left(\sum_{i=1}^{\infty} \sum_{j=1}^{i-1} \frac{1}{2^{i+j}} S_{(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) + \sum_{k_1=1}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{i-1} \frac{1}{2^{2k_1+i+j}} S_{(k_1, k_1)(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) \right. \\ & + \sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{i-1} \frac{1}{2^{2k_1+2k_2+i+j}} S_{(k_1, k_1)(k_2, k_2)(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) \\ & \left. + \sum_{k_1=1}^{\infty} \sum_{k_2=1}^{\infty} \sum_{k_3=1}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{i-1} \frac{1}{2^{2k_1+2k_2+2k_3+i+j}} S_{(k_1, k_1)(k_2, k_2)(k_3, k_3)(i,j)} \left(\frac{1}{2}, \frac{1}{2} \right) + \dots \right) = \left(\frac{7}{10}, \frac{3}{10} \right). \end{aligned}$$

Let $\beta = \{(\frac{3}{10}, \frac{7}{10}), (\frac{7}{10}, \frac{3}{10})\}$. Then, due to symmetry,

$$\int \min_{c \in \beta} \|x - c\|^2 dP = 2 \int_{M((\frac{3}{10}, \frac{7}{10}) | \beta)} \|x - (\frac{3}{10}, \frac{7}{10})\|^2 dP.$$

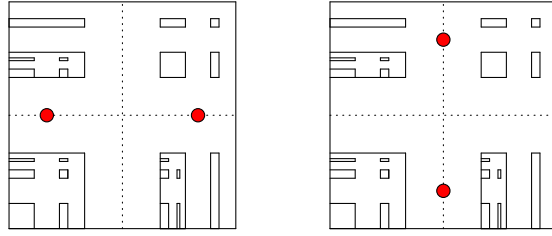


FIGURE 2. Optimal sets of two-means.

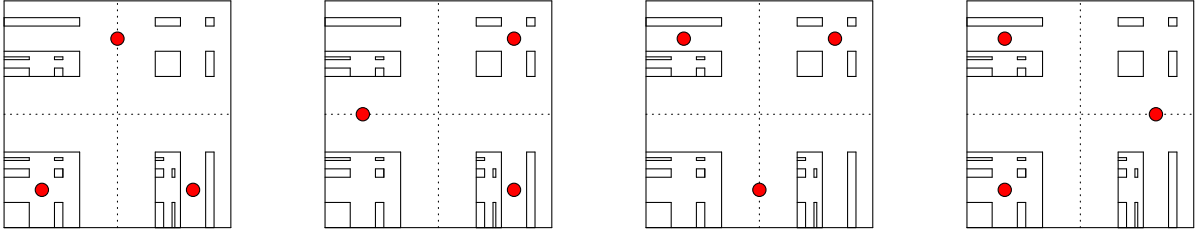
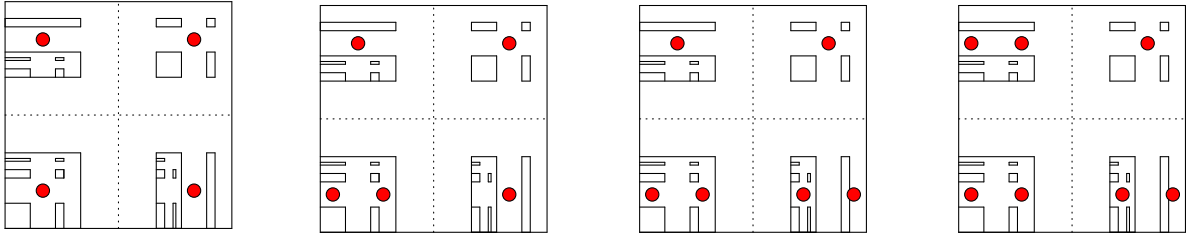


FIGURE 3. Optimal sets of three-means.

FIGURE 4. Optimal sets of n -means for $4 \leq n \leq 7$. Optimal set of 4-means is unique; on the other hand, optimal sets of n -means for $n = 5, 6, 7$ are not unique.

Write

$$\begin{aligned}
 A := & \left(\bigcup_{j=2}^4 J_{(1,1)(1,1)(1,1)(1,1)(1,j)} \right) \cup \left(\bigcup_{j=2}^6 J_{(1,1)(1,1)(1,1)(1,j)} \right) \cup \left(\bigcup_{j=3}^5 J_{((1,1)(1,1)(1,1)(2,j))} \right) \cup \left(\bigcup_{j=2}^8 J_{(1,1)(1,1)(1,j)} \right) \\
 & \cup \left(\bigcup_{j=3}^6 J_{(1,1)(1,1)(2,j)} \right) \cup J_{(1,1)(1,1)(3,4)} \cup \left(\bigcup_{j=2}^8 J_{(1,1)(1,j)} \right) \cup \left(\bigcup_{j=3}^7 J_{(1,1)(2,j)} \right) \cup \left(\bigcup_{j=4}^6 J_{(1,1)(3,j)} \right) \cup \left(\bigcup_{j=2}^{10} J_{(1,j)} \right) \\
 & \cup \left(\bigcup_{j=3}^{10} J_{(2,j)} \right) \cup \left(\bigcup_{j=4}^{10} J_{(3,j)} \right) \cup \left(\bigcup_{j=5}^9 J_{(4,j)} \right) \cup \left(\bigcup_{j=6}^7 J_{(5,j)} \right).
 \end{aligned}$$

Since A is a proper subset of $M((\frac{3}{10}, \frac{7}{10})|\beta)$, we have $\int \min_{c \in \beta} \|x - c\|^2 dP > 2 \int_A \|x - (\frac{3}{10}, \frac{7}{10})\|^2 dP$.

Now using (4), and then upon simplification, it follows that

$$\int \min_{c \in \beta} \|x - c\|^2 dP > 2 \int_A \|x - (\frac{3}{10}, \frac{7}{10})\|^2 dP = 0.13899,$$

which is larger than the distortion error 0.138889 obtained in Lemma 3.1. Hence, the points in an optimal set of two-means can not lie on a oblique line of the affine set. Thus, the assertion of the lemma follows. \square

The following proposition gives the optimal sets of two-means.

Proposition 3.3. *Let P be the affine measure on \mathbb{R}^2 . Then the sets $\{(\frac{1}{6}, \frac{1}{2}), (\frac{5}{6}, \frac{1}{2})\}$ and $\{(\frac{1}{2}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$ form two different optimal sets of two-means with quantization error $\frac{5}{36} = 0.138889$.*

Proof. By Lemma 3.2, it is known that the points in an optimal set of two-means can not lie on an oblique line of the affine set. Thus, by Lemma 3.1, we see that $\{(\frac{1}{6}, \frac{1}{2}), (\frac{5}{6}, \frac{1}{2})\}$ forms an optimal set of two-means with quantization error 0.138889. Due to symmetry, $\{(\frac{1}{2}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$ forms another optimal set of two-means (see Figure 2), and thus the proposition is yielded. \square

The following proposition gives an optimal set of three-means.

Proposition 3.4. *Let P be the affine measure on \mathbb{R}^2 . Then the set $\{(\frac{1}{6}, \frac{1}{6}), (\frac{5}{6}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$ forms an optimal set of three-means with quantization error $\frac{1}{12}$.*

Proof. Let us first consider a three-point set β given by $\beta = \{(\frac{1}{6}, \frac{1}{6}), (\frac{5}{6}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$. Then, using Lemma 2.8 and equation (4), we have

$$\begin{aligned} \int \min_{a \in \beta} \|x - a\|^2 dP &= \int_{J_{(1,1)}} \|x - (\frac{1}{6}, \frac{1}{6})\|^2 dP + \int_{J_{(1,1)(\infty, \emptyset)}} \|x - (\frac{5}{6}, \frac{1}{6})\|^2 dP \\ &\quad + \int_{J_{(1,1)(\emptyset, \infty)} \cup J_{(1,1)(\infty, \infty)}} \|x - (\frac{1}{2}, \frac{5}{6})\|^2 dP = \frac{1}{12} = 0.0833333. \end{aligned}$$

Since V_3 is the quantization error for an optimal set of three-means, we have $\frac{1}{12} \geq V_3$. Let $\alpha = \{(a_i, b_i) : 1 \leq i \leq 3\}$ be an optimal set of three-means. Since the optimal points are the centroids of their own Voronoi regions, we have $\alpha \subset [0, 1] \times [0, 1]$. Let $A_1 = [0, \frac{1}{3}] \times [0, \frac{1}{3}]$, $A_2 = [\frac{2}{3}, 1] \times [0, \frac{1}{3}]$, $A_3 = [0, \frac{1}{3}] \times [\frac{2}{3}, 1]$, and $A_4 = [\frac{2}{3}, 1] \times [\frac{2}{3}, 1]$. Note that the centroids of A_1, A_2, A_3 and A_4 with respect to the probability distribution P are respectively $(\frac{1}{6}, \frac{1}{6}), (\frac{5}{6}, \frac{1}{6}), (\frac{1}{6}, \frac{5}{6})$ and $(\frac{5}{6}, \frac{5}{6})$. Suppose that α does not contain any point from $\bigcup_{i=1}^4 A_i$. Then, we can assume that all the points of α are on the line $x_2 = \frac{1}{2}$, i.e., $\alpha = \{(a_i, \frac{1}{2}) : 1 \leq i \leq 3\}$ with $a_1 < a_2 < a_3$. If $a_1 > \frac{1}{3}$ quantization error can be strictly reduced by moving the point $(a_1, \frac{1}{2})$ to $(\frac{1}{3}, \frac{1}{2})$. So, we can assume that $a_1 \leq \frac{1}{3}$. Similarly, we can show that $a_3 \geq \frac{2}{3}$. Now, if $a_2 < \frac{1}{3}$, then $A_3 \cup A_4 \subset M((a_3, \frac{1}{2})|\alpha)$. Moreover, for any $(x_1, x_2) \in J_{(1,1)(1,1)} \cup J_{(1,1)(1,3)}$, we have $\min_{c \in \alpha} \|(x_1, x_2) - c\|^2 \geq (\frac{7}{18})^2$ and so by (4) and Lemma 2.8, we obtain

$$\begin{aligned} \int \min_{c \in \alpha} \|x - c\|^2 dP &= \int_{J_{(1,1)(1,1)} \cup J_{(1,1)(1,3)}} \min_{c \in \alpha} \|x - c\|^2 dP + \int_{J_{(1,1)(\infty, \emptyset)} \cup J_{(1,1)(\infty, \infty)}} \min_{c \in \alpha} \|x - c\|^2 dP \\ &\geq \frac{1}{16} \left(\left(\frac{1}{81} + \frac{1}{81} \right) \frac{1}{8} + \left(\frac{7}{18} \right)^2 \right) + \frac{1}{16} \left(\left(\frac{1}{9} + \frac{1}{27^2} \right) \frac{1}{8} + \left(\frac{7}{18} \right)^2 \right) + \int_{J_{(1,1)(\infty, \emptyset)} \cup J_{(1,1)(\infty, \infty)}} \|x - (\frac{5}{6}, \frac{1}{2})\|^2 dP \\ &= \frac{1}{16} \left(\left(\frac{1}{81} + \frac{1}{81} \right) \frac{1}{8} + \left(\frac{7}{18} \right)^2 \right) + \frac{1}{16} \left(\left(\frac{1}{9} + \frac{1}{27^2} \right) \frac{1}{8} + \left(\frac{7}{18} \right)^2 \right) + \frac{5}{72} = \frac{1043}{11664} = 0.0894204 > V_3, \end{aligned}$$

which is a contradiction, and so $a_2 \geq \frac{1}{3}$ must be true. If $a_2 > \frac{2}{3}$, similarly we can show contradiction arises. So, $\frac{1}{3} < a_2 < \frac{2}{3}$. Next, suppose that $\frac{1}{2} \leq a_2 < \frac{2}{3}$. Then, we have $\frac{1}{2}(a_1 + a_2) \leq \frac{1}{3}$ which implies $a_1 \leq \frac{1}{6}$, for otherwise quantization error can be strictly reduced by moving a_2 to $(\frac{2}{3}, \frac{1}{2})$, contradicting the fact that α is an optimal set. Then, $\bigcup_{j=1}^{\infty} J_{(1,1)(1,j)} \cup$

$\bigcup_{i=2,j=1}^{\infty} J_{(1,i)(1,j)} \subset M((a_1, \frac{1}{2})|\alpha)$ and

$$E(X : X \in \bigcup_{j=1}^{\infty} J_{(1,1)(1,j)} \cup \bigcup_{i=2,j=1}^{\infty} J_{(1,i)(1,j)}) = \left(\frac{1}{18}, \frac{1}{2} \right).$$

Since for any $(x_1, x_2) \in \bigcup_{i=2, j=1}^{\infty} J_{(1,1)(i,j)} \cup \bigcup_{k=1, i=2, j=1}^{\infty} J_{(k,2)(i,j)}$, we have $\min_{c \in \alpha} \|(x_1, x_2) - c\|^2 \geq \|(x_1, x_2) - (\frac{1}{6}, \frac{1}{2})\|^2$. Thus, writing

$$A := \bigcup_{j=1}^{\infty} J_{(1,1)(1,j)} \cup \bigcup_{i=2, j=1}^{\infty} J_{(1,i)(1,j)} \text{ and } B = \bigcup_{i=2, j=1}^{\infty} J_{(1,1)(i,j)} \cup \bigcup_{k=1, i=2, j=1}^{\infty} J_{(k,2)(i,j)}, \text{ we have}$$

$$\begin{aligned} & \int \min_{c \in \alpha} \|x - c\|^2 dP > \int_A \|(x_1, x_2) - (\frac{1}{18}, \frac{1}{2})\|^2 dP + \int_B \|(x_1, x_2) - (\frac{1}{6}, \frac{1}{2})\|^2 dP \\ &= 2 \int_{\bigcup_{j=1}^{\infty} J_{(1,1)(1,j)}} \|x - (\frac{1}{18}, \frac{1}{2})\|^2 dP + \int_{\bigcup_{i=2, j=1}^{\infty} J_{(1,1)(i,j)}} \|x - (\frac{1}{6}, \frac{1}{2})\|^2 dP \\ & \quad + \int_{\bigcup_{k=1, i=2, j=1}^{\infty} J_{(k,2)(i,j)}} \|x - (\frac{1}{6}, \frac{1}{2})\|^2 dP \\ &= 2 \cdot \frac{41}{2592} + \frac{5}{288} + \frac{551}{14688} = \frac{953}{11016} = 0.0865105 > V_3, \end{aligned}$$

which is a contradiction. Similarly, if we assume $\frac{1}{3} \leq a_2 < \frac{1}{2}$, a contradiction will arise. Therefore, all the points in α can not lie on the line $x_2 = \frac{1}{2}$. Let (a_1, b_1) and (a_3, b_3) lie on the line $x_2 = \frac{1}{2}$, and (a_2, b_2) is above or below the horizontal line $x_2 = \frac{1}{2}$. If (a_2, b_2) is above the horizontal line then the quantization error can be strictly reduced by moving (a_1, b_1) to A_1 and (a_3, b_3) to A_2 contradicting the fact that α is an optimal set. Similarly, if (a_2, b_2) is below the horizontal line a contradiction will arise. All these contradictions arise due to our assumption that α does not contain any point from $\bigcup_{i=1}^4 A_i$. Hence, α contains at least one point from $\bigcup_{i=1}^4 A_i$. In order to complete the proof of the Proposition, first we will prove the following claim:

Claim. $\text{card}(\{i : \alpha \cap A_i \neq \emptyset, 1 \leq i \leq 4\}) = 2$.

For the sake of contradiction, assume that $\text{card}(\{i : \alpha \cap A_i \neq \emptyset, 1 \leq i \leq 4\}) = 1$. Then, without any loss of generality we assume that $(a_1, b_1) \in A_1$ and $(a_i, b_i) \notin A_2 \cup A_3 \cup A_4$ for $i = 2, 3$. Due to symmetry of the affine set with respect to the diagonal $x_2 = x_1$, we can assume that $(a_1, b_1) \in A_1$ lies on the diagonal $x_2 = x_1$; (a_2, b_2) and (a_3, b_3) are equidistant from the diagonal $x_2 = x_1$ and are in opposite sides of the diagonal $x_2 = x_1$. Let us now consider the following cases:

Case 1. Assume that both (a_2, b_2) and (a_3, b_3) are below the diagonal $x_2 = 1 - x_1$, but not in $A_1 \cup A_2 \cup A_3$. Let (a_2, b_2) be above the diagonal $x_2 = x_1$ and (a_3, b_3) be below the diagonal $x_2 = x_1$. In that case, quantization error can be strictly reduced by moving (a_2, b_2) to A_3 and (a_3, b_3) to A_2 which contradicts the optimality of α .

Case 2. Assume that both (a_2, b_2) and (a_3, b_3) are above the diagonal $x_2 = 1 - x_1$. Let (a_2, b_2) lie above the diagonal $x_2 = x_1$ and (a_3, b_3) lie below the diagonal $x_2 = x_1$. Then, due to symmetry we can assume that $(a_1, b_1) = (\frac{1}{6}, \frac{1}{6})$ which is the centroid of A_1 , $(a_2, b_2) = (\frac{1}{2}, \frac{5}{6})$ which is the midpoint of the line segment joining the centroids of A_3 and A_4 , $(a_3, b_3) = (\frac{5}{6}, \frac{1}{2})$ which is the midpoint of the line segment joining the centroids of A_2 and A_4 . Then,

$$\begin{aligned}
\int \min_{c \in \alpha} \|x - c\|^2 dP &= \int_{J_{(1,1)}} \min_{c \in \alpha} \|x - c\|^2 dP + \int_{J_{(1,1)(\emptyset, \infty)}} \min_{c \in \alpha} \|x - c\|^2 dP \\
&\quad + \int_{J_{(1,1)(\infty, \emptyset)}} \min_{c \in \alpha} \|x - c\|^2 dP + \int_{J_{(1,1)(\infty, \infty)}} \min_{c \in \alpha} \|x - c\|^2 dP \\
&\geq \frac{1}{4} \left(\frac{1}{9} + \frac{1}{9} \right) \frac{1}{8} + \int_{J_{(1,1)(\emptyset, \infty)}} \|x - (\frac{1}{2}, \frac{5}{6})\|^2 dP + \int_{J_{(1,1)(\infty, \emptyset)}} \|x - (\frac{5}{6}, \frac{1}{2})\|^2 dP \\
&\quad + \int_{\bigcup_{i=2}^{\infty} \bigcup_{j=i+1}^{\infty} J_{(i,j)}} \|x - (\frac{1}{2}, \frac{5}{6})\|^2 dP \\
&= \frac{1}{4} \left(\frac{1}{9} + \frac{1}{9} \right) \frac{1}{8} + \frac{5}{144} + \frac{5}{144} + \frac{1381}{166320} = \frac{7043}{83160} = 0.0846922 > V_3,
\end{aligned}$$

which is a contradiction. Thus, under the assumption $\text{card}(\{i : \alpha \cap A_i \neq \emptyset, 1 \leq i \leq 4\}) = 1$, we arrive at a contradiction.

Next, for the sake of contradiction, assume that $\text{card}(\{i : \alpha \cap A_i \neq \emptyset, 1 \leq i \leq 4\}) = 3$. Then, without any loss of generality we assume that $(a_1, b_1) \in A_3$, $(a_2, b_2) \in A_2$ and $(a_3, b_3) \in A_4$. Let A_{11} and A_{12} be the regions of A_1 which are respectively above and below the diagonal of A_1 passing through $(0, 0)$. Due to symmetry, we must have $A_3 \cup A_{11} \subset M((a_1, b_1)|\alpha)$ and $A_2 \cup A_{12} \subset M((a_2, b_2)|\alpha)$. Notice that $A_3 \cup A_{11} \subset M((a_1, b_1)|\alpha)$ implies

$$A_3 \cup \bigcup_{i=1, j=i+1} J_{(1,1)(i,j)} \cup \bigcup_{k=1, i=1 \atop j=i+1} J_{(1,1)(k,k)(i,j)} \subset M((a_1, b_1)|\alpha),$$

and using (1), we have

$$E(X : X \in A_3 \cup \bigcup_{i=1, j=i+1} J_{(1,1)(i,j)} \cup \bigcup_{k=1, i=1 \atop j=i+1} J_{(1,1)(k,k)(i,j)}) = \left(\frac{1385}{9438}, \frac{6173}{9438} \right) = (0.146747, 0.654058),$$

which shows that the point (a_1, b_1) falls below the line $x_2 = \frac{2}{3}$, which is a contradiction as we assumed that $(a_1, b_1) \in A_3$. This contradiction arises due to our assumption that $\text{card}(\{i : \alpha \cap A_i \neq \emptyset, 1 \leq i \leq 4\}) = 3$. Hence, we conclude that $\text{card}(\{i : \alpha \cap A_i \neq \emptyset, 1 \leq i \leq 4\}) = 2$, which proves the claim.

By the claim, we assume that $(a_1, b_1) \in A_1$ and $(a_3, b_3) \in A_2$. Notice that A_1, A_2, A_3, A_4 are geometrically symmetric as well as their corresponding centroids are symmetrically distributed over the square $[0, 1] \times [0, 1]$. Without any loss of generality, we can assume that the optimal point (a_1, b_1) is the centroid of A_1 , i.e., $(a_1, b_1) = (\frac{1}{6}, \frac{1}{6})$. Then, due to symmetry with respect to the line $x_1 = \frac{1}{2}$, it follows that $(a_3, b_3) = \text{centroid of } A_2 = (\frac{5}{6}, \frac{1}{6})$, and (a_2, b_2) lies on $x_1 = \frac{1}{2}$ but above the line $x_2 = \frac{1}{2}$. Now, notice that

$$\min_{(a_3, b_3) \in [\frac{1}{3}, \frac{2}{3}] \times [\frac{2}{3}, 1]} \{ \|(\frac{1}{6}, \frac{5}{6}) - (a_3, b_3)\|^2 + \|(\frac{5}{6}, \frac{5}{6}) - (a_3, b_3)\|^2 \} = \frac{2}{9},$$

which occurs when $(a_3, b_3) = \text{center of } [\frac{1}{3}, \frac{2}{3}] \times [\frac{2}{3}, 1] = (\frac{1}{2}, \frac{5}{6})$. Moreover, the three points $(\frac{1}{6}, \frac{1}{6})$, $(\frac{5}{6}, \frac{1}{6})$ and $(\frac{1}{2}, \frac{5}{6})$ are the centroids of their own Voronoi regions. Thus, $\{(\frac{1}{6}, \frac{1}{6}), (\frac{5}{6}, \frac{1}{6}), (\frac{1}{2}, \frac{5}{6})\}$ forms an optimal set of three-means with quantization error $V_3 = \frac{1}{12} = 0.0833333$. Hence, the proposition follows. \square

Remark 3.5. Due to symmetry, in addition to the optimal set given in Proposition 3.4, there are three more optimal sets of three-means with quantization error $V_3 = \frac{1}{12}$ (see Figure 3).

4. AFFINE MEASURES

In this section, first we give some basic preliminaries and show that affine measure under consideration is the direct product of the Cantor distribution.

Let P_c be the Cantor distribution generated by the two similitudes U_1 and U_2 such that $U_1(x) = \frac{1}{3}x$ and $U_2(x) = \frac{1}{3}x + \frac{2}{3}$ for all $x \in \mathbb{R}$, i.e., $P_c = \frac{1}{2}P_c \circ U_1^{-1} + \frac{1}{2}P_c \circ U_2^{-1}$. Then P_c has support the Cantor set C generated by the similitudes U_1 and U_2 . By a word σ of length k over the alphabet $\{1, 2\}$, it is meant $\sigma := \sigma_1\sigma_2 \cdots \sigma_k \in \{1, 2\}^k$, $k \geq 1$. By a word of length zero it is meant the empty word \emptyset . $\{1, 2\}^*$ represents the set of all words over the alphabet $\{1, 2\}$ including the empty word \emptyset . Length of a word $\sigma \in \{1, 2\}^*$ is denoted by $|\sigma|$. If $\sigma = \sigma_1\sigma_2 \cdots \sigma_k$, we write $U_\sigma := U_{\sigma_1} \circ U_{\sigma_2} \circ \cdots \circ U_{\sigma_k}$. U_\emptyset represents the identity mapping on \mathbb{R} . By u_σ we represent the similarity ratio of U_σ . If X_c is the random variable with distribution P_c , then $E(X_c) = \frac{1}{2}$ and $V(X_c) = \text{Variance of } X_c = \frac{1}{8}$ (see [8]). For $\sigma \in \{1, 2\}^*$, write $A(\sigma) := U_\sigma(\frac{1}{2})$. Notice that for $\sigma \in \{1, 2\}^*$, we have $\frac{1}{2}(A(\sigma 1) + A(\sigma 2)) = A(\sigma)$, $u_\sigma = \frac{1}{3^{|\sigma|}}$, and for the empty word \emptyset , $A(\emptyset) = \frac{1}{2}$. For $\sigma \in \{1, 2\}^*$ define $A_\sigma := U_\sigma[0, 1]$. For any positive integer n , by 2^{*n} it is meant the concatenation of the symbol 2 with itself n -times successively, i.e., $2^{*n} = 222 \cdots (n \text{ times})$, with the convention that 2^{*0} is the empty word. For any positive integer k , by $\{1, 2\}^{k*2}$ it is meant the direct product of the set $\{1, 2\}^k$ with itself. By $\{1, 2\}^{0*2}$ it is meant the set $\{(\emptyset, \emptyset)\}$. Also, recall the notations defined in Section 2. Let us now introduce the map $F : \mathbb{N}^* \cup \{(\sigma, \infty) : \sigma \in \mathbb{N}^*\} \rightarrow \{1, 2\}^*$ such that

$$(6) \quad F(x) = \begin{cases} f(\sigma_1)f(\sigma_2) \cdots f(\sigma_{|\sigma|}) & \text{if } x = \sigma = \sigma_1\sigma_2 \cdots \sigma_{|\sigma|}, \\ f(\sigma_1)f(\sigma_2) \cdots f(\sigma_{|\sigma|}, \infty) & \text{if } x = (\sigma_1\sigma_2 \cdots \sigma_{|\sigma|}, \infty), \\ \emptyset & \text{if } x = \emptyset, \end{cases}$$

where $f : \mathbb{N} \cup \{(n, \infty) : n \in \mathbb{N}\} \rightarrow \{1, 2\}^* \setminus \{\emptyset\}$ is such that

$$f(x) = \begin{cases} 2^{*(n-1)}1 & \text{if } x = n \text{ for some } n \in \mathbb{N}, \\ 2^{*n} & \text{if } x = (n, \infty) \text{ for some } n \in \mathbb{N}. \end{cases}$$

It is easy to see that the function f is one-to-one and onto which yields the fact that F is also one-to-one and onto. For any $\sigma \in \mathbb{N}^*$, write $AF(\sigma) := A(F(\sigma))$ and $AF(\sigma, \infty) := A(F(\sigma, \infty))$.

Remark 4.1. In the sequel, we will show that the map F is useful to convert the infinitely generated affine measure P to a finitely generated affine measure $P_c \times P_c$.

Lemma 4.2. Let T_k for $k \geq 1$ be the infinite collection of similitudes as defined in Remark 2.3, and U_1, U_2 be the two similitudes generating the Cantor set. Then, for any $\sigma \in \mathbb{N}^*$ and $x \in \mathbb{R}$, we have $T_\sigma(x) = U_{F(\sigma)}(x)$.

Proof. Take any $x \in \mathbb{R}$. If $\sigma = 1$, then $T_1(x) = \frac{1}{3}x = U_1(x) = U_{F(1)}(x)$. Let us assume that the lemma is true if $\sigma = k$ for some positive integer k , i.e., $T_k(x) = U_{F(k)}(x)$. We now show that $T_{k+1}(x) = U_{F(k+1)}(x)$. See that

$$\begin{aligned} U_{F(k+1)}(x) &= U_{2^{*k}1}(x) = U_{2^{*(k-1)}21}(x) = U_{2^{*(k-1)}U_{21}}(x) = U_{2^{*(k-1)}}(\frac{1}{9}x + \frac{2}{9}) \\ &= U_{2^{*(k-1)}1}(3(\frac{1}{9}x + \frac{2}{9})) = U_{F(k)}(\frac{1}{3}x + \frac{2}{3}) = T_k(\frac{1}{3}x + \frac{2}{3}) = \frac{1}{3^k}(\frac{1}{3}x + \frac{2}{3}) + 1 - \frac{1}{3^{k-1}} \\ &= \frac{1}{3^{k+1}}x + 1 - \frac{1}{3^k} = T_{k+1}(x). \end{aligned}$$

Thus, by the Principle of Mathematical Induction, one can say that $T_k(x) = U_{F(k)}(x)$ for all $k \in \mathbb{N}$. Again, for any $\tau, \delta \in \mathbb{N}^*$, by (6), it follows that $F(\sigma\delta) = F(\sigma)F(\delta)$. Hence, for any $\sigma = \sigma_1\sigma_2 \cdots \sigma_n \in \mathbb{N}^*$, $n \geq 1$, we have

$$T_\sigma(x) = T_{\sigma_1} \circ T_{\sigma_2} \circ \cdots \circ T_{\sigma_n}(x) = U_{F(\sigma_1)} \circ U_{F(\sigma_2)} \circ \cdots \circ U_{F(\sigma_n)}(x) = U_{F(\sigma)}(x),$$

which completes the proof of the lemma. \square

Lemma 4.3. *Let $\omega \in I^*$, and F be the function as defined in (6). Then for $r = 1, 2$, we have $AF(\omega^{(r)}) = S_{\omega}^{(r)}(\frac{1}{2})$, and $AF(\omega^{(r)}, \infty) = S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(r)}(\frac{1}{2}) + s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(r)}$.*

Proof. Using Lemma 4.2, we have

$$AF(\omega^{(1)}) = U_{F(\omega^{(1)})}(\frac{1}{2}) = T_{\omega^{(1)}}(\frac{1}{2}) = S_{\omega}^{(1)}(\frac{1}{2}), \text{ and similarly } AF(\omega^{(2)}) = S_{\omega}^{(2)}(\frac{1}{2}).$$

Without any loss of generality, we assume $\omega = (i_1, j_1)(i_2, j_2) \cdots (i_k, j_k)$ for $k \geq 1$. Then,

$$\begin{aligned} AF(\omega^{(1)}, \infty) &= U_{F(i_1 i_2 \cdots i_k, \infty)}(\frac{1}{2}) = U_{F(i_1 i_2 \cdots i_{k-1})} \circ U_{F(i_k, \infty)}(\frac{1}{2}) = U_{F(i_1 i_2 \cdots i_{k-1})} \circ U_{2^{*i_k}}(\frac{1}{2}) \\ &= U_{F(i_1 i_2 \cdots i_{k-1})} \circ U_{2^{*i_k}}(U_1^{-1}(\frac{1}{2})) = U_{F(i_1 i_2 \cdots i_{k-1})} \circ U_{F(i_k+1)}(\frac{3}{2}) = U_{F(i_1 i_2 \cdots i_{k-1}(i_k+1))}(\frac{3}{2}) \\ &= T_{i_1 i_2 \cdots i_{k-1}(i_k+1)}(\frac{3}{2}) = S_{\omega^-(i_k+1, j_k+1)}^{(1)}(\frac{3}{2}). \end{aligned}$$

Since, $S_{(i_k+1, j_k+1)}^{(1)}(\frac{3}{2}) - S_{(i_k+1, j_k+1)}^{(1)}(\frac{1}{2}) = \frac{1}{3^{i_k+1}} \frac{3}{2} + 1 - \frac{1}{3^{i_k}} - \frac{1}{3^{i_k+1}} \frac{1}{2} - 1 + \frac{1}{3^{i_k}} = \frac{1}{3^{i_k+1}}$, we have

$$\begin{aligned} S_{\omega^-(i_k+1, j_k+1)}^{(1)}(\frac{3}{2}) - S_{\omega^-(i_k+1, j_k+1)}^{(1)}(\frac{1}{2}) &= s_{\omega^-(i_k+1, j_k+1)}^{(1)}(S_{(i_k+1, j_k+1)}^{(1)}(\frac{3}{2}) - S_{(i_k+1, j_k+1)}^{(1)}(\frac{1}{2})) = s_{\omega^-(i_k+1, j_k+1)}^{(1)} \frac{1}{3^{i_k+1}} \\ &= s_{\omega^-(i_k+1, j_k+1)}^{(1)} = s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}, \end{aligned}$$

yielding $AF(\omega^{(1)}, \infty) = S_{\omega^-(i_k+1, j_k+1)}^{(1)}(\frac{3}{2}) = S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}(\frac{1}{2}) + s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(1)}$, and similarly, $AF(\omega^{(2)}, \infty) = S_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}(\frac{1}{2}) + s_{\omega^-(\omega_{|\omega|}^{(1)}+1, \omega_{|\omega|}^{(2)}+1)}^{(2)}$. Thus, the proof of the lemma is complete. \square

Remark 4.4. By Lemma 2.6 and Lemma 4.3 for any $\omega \in I^*$, we have

$$\begin{aligned} a(\omega) &= (AF(\omega^{(1)}), AF(\omega^{(2)})), \quad a(\omega(\infty, \infty)) = (AF(\omega^{(1)}, \infty), AF(\omega^{(2)}, \infty)), \\ a(\omega(\infty, \emptyset)) &= (AF(\omega^{(1)}, \infty), AF(\omega^{(2)})), \quad a(\omega(\emptyset, \infty)) = (AF(\omega^{(1)}), AF(\omega^{(2)}, \infty)). \end{aligned}$$

The following example illustrates the outcome of the lemma above.

Example 4.5. $a((1, 1)) = (AF(1), AF(1)) = (A(1), A(1)) = (\frac{1}{6}, \frac{1}{6})$,
 $a((1, 1)(\infty, \emptyset)) = (AF(1, \infty), AF(1)) = (A(2), A(1)) = (\frac{5}{6}, \frac{1}{6})$,
 $a((1, 1)(\emptyset, \infty)) = (AF(1), AF(1, \infty)) = (A(1), A(2)) = (\frac{1}{6}, \frac{5}{6})$,
 $a((1, 1)(\infty, \infty)) = (AF(1, \infty), AF(1, \infty)) = (A(2), A(2)) = (\frac{5}{6}, \frac{5}{6})$,
 $a((1, 1)(1, 1)) = (AF(11), AF(11)) = (A(11), A(11)) = (\frac{1}{18}, \frac{1}{18})$,
 $a((1, 1)(1, 1)(\infty, \emptyset)) = (AF(11, \infty), AF(11)) = (A(12), A(11)) = (\frac{5}{18}, \frac{1}{18})$,
 $a((1, 1)(1, 1)(\emptyset, \infty)) = (AF(11), AF(11, \infty)) = (A(11), A(12)) = (\frac{1}{18}, \frac{5}{18})$, and
 $a((1, 1)(1, 1)(\infty, \infty)) = (AF(11, \infty), AF(11, \infty)) = (A(12), A(12)) = (\frac{5}{18}, \frac{5}{18})$, etc.

Lemma 4.6. *Let $\mu = \sum_{k=1}^{\infty} \frac{1}{2^k} \mu \circ T_k^{-1}$. Then, for any $\sigma \in \mathbb{N}^*$, we have $\mu(T_{\sigma}[0, 1]) = P_c(A_{F(\sigma)})$, where $P_c := \frac{1}{2}P_c \circ U_1^{-1} + \frac{1}{2}P_c \circ U_2^{-1}$.*

Proof. Without any loss of generality, let $\sigma = i_1 i_2 \cdots i_k$ for any $k \geq 1$. See that $F(\sigma) = F(i_1)F(i_2) \cdots F(i_k)$, and thus $|F(\sigma)| = |F(i_1)| + |F(i_2)| + \cdots + |F(i_k)| = i_1 + i_2 + \cdots + i_k$, which implies

$$\mu(T_{\sigma}[0, 1]) = \frac{1}{2^{i_1+i_2+\cdots+i_k}} = \frac{1}{2^{|F(\sigma)|}} = P_c(A_{F(\sigma)}),$$

which is the lemma. \square

The following proposition plays an important role in the paper.

Proposition 4.7. *Let P be the affine measure. Then, $P = P_c \times P_c$, where P_c is the Cantor distribution.*

Proof. Borel σ -algebra on the affine set is generated by all sets of the form $J_{(\delta, \tau)}$ for $(\delta, \tau) \in I^*$, where $J_{(\delta, \tau)} = S_{(\delta, \tau)}([0, 1] \times [0, 1])$. Notice that

$$J_{(\delta, \tau)} = T_\delta[0, 1] \times T_\tau[0, 1] = U_{F(\delta)}[0, 1] \times U_{F(\tau)}[0, 1] = A_{F(\delta)} \times A_{F(\tau)}.$$

Again, the sets of the form A_α , where $\alpha \in \{1, 2\}^*$, generate the Borel σ -algebra on the Cantor set C . Thus, we see that Borel σ -algebra of the affine set is same as the product of the Borel σ -algebras on the Cantor set. Moreover, for any $(\delta, \tau) \in I^*$, by Remark 2.3 and Lemma 4.6, we have

$$P(J_{(\delta, \tau)}) = \mu(T_\delta[0, 1])\mu(T_\tau[0, 1]) = P_c(A_{F(\delta)})P_c(A_{F(\tau)}) = (P_c \times P_c)(A_{F(\delta)} \times A_{F(\tau)}).$$

Hence, the proposition follows. \square

Remark 4.8. By Proposition 4.7, it follows that the optimal sets of n -means for P are same as the optimal sets n -means for the product measure $P_c \times P_c$. Moreover, for $k \geq 1$ we can write

$$P = P_c \times P_c = \sum_{(\sigma, \tau) \in \{1, 2\}^{k*2}} \frac{1}{4^k} (P_c \times P_c) \circ (U_\sigma, U_\tau)^{-1},$$

where for $(x_1, x_2) \in \mathbb{R}^2$, $(U_\sigma, U_\tau)^{-1}(x_1, x_2) = (U_\sigma^{-1}(x_1), U_\tau^{-1}(x_2))$.

5. OPTIMAL SETS OF n -MEANS FOR ALL $n \geq 4$

In this section we will give closed formulas to determine the optimal sets of n -means and the n th quantization error for all $n \geq 4$. For $(\sigma, \tau) \in \{1, 2\}^{k*2}$, write $A_{(\sigma, \tau)} := A_\sigma \times A_\tau$ and $U_{(\sigma, \tau)} := (U_\sigma, U_\tau)$.

Lemma 5.1. *Let α be an optimal set of n -means with $n \geq 4$. Then, $\alpha \cap A_{(i, j)} \neq \emptyset$ for all $1 \leq i, j \leq 2$.*

Proof. Let α be an optimal set of n -means for $n \geq 4$. As the optimal points are the centroids of their own Voronoi regions we have $\alpha \subset A_\emptyset \times A_\emptyset := [0, 1] \times [0, 1]$.

Consider the four-point set β given by $\beta = \{(A(i), A(j)) : 1 \leq i, j \leq 2\}$. Then,

$$\int \min_{c \in \beta} \|x - c\|^2 dP = \sum_{i, j=1}^2 \int_{A_{(i, j)}} \|x - (A(i), A(j))\|^2 d(P_c \times P_c) = \sum_{i, j=1}^2 \frac{1}{4} \left(\frac{1}{9} + \frac{1}{9} \right) \frac{1}{8} = \frac{1}{36}.$$

Since V_4 is the quantization error of four-means, we have $\frac{1}{36} = 0.0277778 \geq V_4 \geq V_n$.

Now, for the sake of contradiction, assume that α does not contain any point from $\bigcup_{i, j=1}^2 A_{(i, j)}$. We know that

$$(7) \quad \sum_{(a, b) \in \alpha} (a, b) P(M(a, b) | \alpha) = \left(\frac{1}{2}, \frac{1}{2} \right).$$

If all the points of α are below the line $x_2 = \frac{1}{2}$, i.e., if $b < \frac{1}{2}$ then by (7), we see that $\frac{1}{2} = \sum_{(a, b) \in \alpha} b P(M(a, b) | \alpha) < \sum_{(a, b) \in \alpha} \frac{1}{2} P(M(a, b) | \alpha) = \frac{1}{2}$, which is a contradiction. Similarly, it follows that if all the points of α are above the line $x_2 = \frac{1}{2}$, or left of the line $x_1 = \frac{1}{2}$, or right of the line $x_1 = \frac{1}{2}$, a contradiction will arise.

Next, suppose that all the points of α are on the line $x_2 = \frac{1}{2}$. We will consider two cases: $n = 4$ and $n > 4$. When $n = 4$, let $\alpha = \{(a_i, \frac{1}{2}) : 1 \leq i \leq 4\}$ with $a_i < a_j$ for $i < j$. Due to symmetry, we can assume that boundary of the Voronoi regions of the points $(a_1, \frac{1}{2})$, $(a_2, \frac{1}{2})$, $(a_3, \frac{1}{2})$, and

$(a_4, \frac{1}{2})$ are respectively $x_1 = \frac{1}{6}$, $x_1 = \frac{1}{2}$, and $x_1 = \frac{5}{6}$ yielding $\alpha = \{(\frac{1}{18}, \frac{1}{2}), (\frac{5}{18}, \frac{1}{2}), (\frac{13}{18}, \frac{1}{2}), (\frac{17}{18}, \frac{1}{2})\}$, and then writing $B := A_{(11,11)} \cup A_{(11,12)} \cup A_{(11,21)} \cup A_{(11,22)}$, by symmetry we have

$$\begin{aligned} \int \min_{c \in \alpha} \|x - c\|^2 dP &= 4 \int_B \|x - (\frac{1}{18}, \frac{1}{2})\|^2 d(P_c \times P_c) \\ &= 8 \int_{A_{(11,11)}} \|x - (\frac{1}{18}, \frac{1}{2})\|^2 d(P_c \times P_c) + 8 \int_{A_{(11,12)}} \|x - (\frac{1}{18}, \frac{1}{2})\|^2 d(P_c \times P_c) \\ &= 8(\frac{65}{5184} + \frac{17}{5184}) = \frac{41}{324} = 0.126543 > V_4, \end{aligned}$$

which is a contradiction. We consider the case $n > 4$. Since for any $(x_1, x_2) \in \bigcup_{i,j=1}^2 A_{ij}$, $\min_{c \in \alpha} \|(x_1, x_2) - c\|^2 \geq \frac{1}{36}$, we have

$$\int \min_{c \in \alpha} \|x - c\|^2 dP = \sum_{i,j=1}^2 \int_{A_{(i,j)}} \min_{c \in \alpha} \|x - c\|^2 d(P_c \times P_c) \geq \sum_{i,j=1}^2 \int_{A_{(i,j)}} \frac{1}{36} d(P_c \times P_c) = \frac{1}{36},$$

which implies $\frac{1}{36} \geq V_4 > V_n$, a contradiction. Thus, we see that all the points of α can not lie on $x_2 = \frac{1}{2}$. Similarly, all the points of α can not lie on $x_1 = \frac{1}{2}$.

Notice that the lines $x_1 = \frac{1}{2}$ and $x_2 = \frac{1}{2}$ partition the square $[0, 1] \times [0, 1]$ into four quadrants with center $(\frac{1}{2}, \frac{1}{2})$. If $n = 4k$ for some positive integer k , due to symmetry, we can assume that each quadrant contains k -points from the set α . But then, any of the k points in the quadrant containing a basic rectangle $A_{(i,j)}$ can be moved to $A_{(i,j)}$ which strictly reduce the quantization error, and it gives a contradiction as we assumed that the set α is an optimal set of n -means and α does not contain any point from $A_{(i,j)}$ for $1 \leq i, j \leq 2$.

If $n = 4k + 1, 4k + 2$, or $n = 4k + 3$, then, again due to symmetry, each quadrant gets at least k points. Then, as in the case $n = 4k$, here also one can strictly reduce the quantization error by moving a point in the quadrant containing a basic rectangle $A_{(i,j)}$ to $A_{(i,j)}$ for $1 \leq i, j \leq 2$, which is a contradiction.

Thus, we have proved that $\alpha \cap A_{(i,j)} \neq \emptyset$ for all $1 \leq i, j \leq 2$. □

Lemma 5.2. *Let α be an optimal set of n -means with $n \geq 4$. Then, $\alpha \subset \bigcup_{i,j=1}^2 A_{(i,j)}$.*

Proof. By Lemma 5.1, we know that $\alpha \cap A_{(i,j)} \neq \emptyset$ for all $1 \leq i, j \leq 2$. Now, we will prove the statement by considering four distinct cases:

Case 1: $n = 4k$ for some positive integer $k \geq 1$.

In this case, due to symmetry, we can assume that α contains k points from each of $A_{(i,j)}$, otherwise, quantization error can be reduced by redistributing the points of α equally among $A_{(i,j)}$ for $1 \leq i, j \leq 2$, and so $\alpha \subset \bigcup_{i,j=1}^2 A_{(i,j)}$.

Case 2: $n = 4k + 1$ for some positive integer $k \geq 1$.

In this case, again due to symmetry, we can assume that α contains k points from each of $A_{(i,j)}$, and if possible, one point, say (a, b) , from $A_{(\emptyset, \emptyset)} \setminus \bigcup_{i,j=1}^2 A_{(i,j)}$. By symmetry, one can assume that (a, b) is the midpoint of the line segment joining any two centroids of the basic rectangles $A_{(i,j)}$ for $1 \leq i, j \leq 2$. Let us first take $(a, b) = (\frac{1}{2}, \frac{1}{2})$ which is the center of the affine set. For simplicity, we first assume $k = 1$, i.e., $n = 5$. Then, α contains only one point from each of $A_{(i,j)}$. Let (a_1, b_1) be the point that α takes from $A_{(1,1)}$. As $(\frac{1}{2}, \frac{1}{2})$ lies on the diagonal $x_2 = x_1$, due to symmetry we can also assume that (a_1, b_1) lies on the diagonal $x_2 = x_1$. By Proposition 1.1, we have $P(M((\frac{1}{2}, \frac{1}{2})|\alpha)) > 0$. This yields that $\frac{1}{2}((a_1, b_1) + (\frac{1}{2}, \frac{1}{2})) < (\frac{1}{3}, \frac{1}{3})$

which implies $a_1 < \frac{1}{6}$ and $b_1 < \frac{1}{6}$. Then, we see that

$$\frac{1}{36} = V_4 \approx V_5 = 4 \int_{A_{(1,1)}} \min_{c \in \{(a_1, b_1), (\frac{1}{2}, \frac{1}{2})\}} \|x - c\|^2 dP > \int \min_{c \in \beta} \|x - c\|^2 dP = \frac{2}{81} \geq V_5,$$

where $\beta = \{(\frac{1}{18}, \frac{1}{18}), (\frac{1}{18}, \frac{5}{18}), (\frac{5}{6}, \frac{1}{6}), (\frac{1}{6}, \frac{5}{6}), (\frac{5}{6}, \frac{5}{6})\}$, which is a contradiction. Similarly, if we take (a, b) as the midpoint of a line segments joining the centroids of any two adjacent basic rectangles $A_{(i,j)}$ for $1 \leq i, j \leq 2$, contradiction arises. Proceeding in the similar way, by taking $k = 2, 3, \dots$, we see that contradiction arises at each value k takes. Therefore, $\alpha \subset \bigcup_{i,j=1}^2 A_{(i,j)}$.

Case 3: $n = 4k + 2$ for some positive integer $k \geq 1$.

In this case, due to symmetry, we can assume that α contains k points from each of $A_{(i,j)}$, and if possible, two points, say (a_1, b_1) and (a_2, b_2) , from $A_{(\emptyset, \emptyset)} \setminus \bigcup_{i,j=1}^2 A_{(i,j)}$. Then, by symmetry, we can assume that (a_1, b_1) lies on the midpoint of the line segment joining the centroids of $A_{(1,1)}$, $A_{(2,1)}$; and (a_2, b_2) lies on the midpoint of the line segment joining the centroids of $A_{(1,2)}$ and $A_{(2,2)}$. As in Case 2, this leads to a contradiction. Thus, $\alpha \subset \bigcup_{i,j=1}^2 A_{(i,j)}$.

Case 4: $n = 4k + 3$ for some positive integer $k \geq 1$. Due to symmetry, in this case, we can assume that each of $A_{(1,1)}$ and $A_{(2,1)}$ gets $k + 1$ points; each of $A_{(1,2)}$ and $A_{(2,2)}$ gets k points. The remaining one point lies on the midpoint of the line segment joining the centroids of $A_{(1,2)}$ and $A_{(2,2)}$. But, in that case, proceeding as in Case 2, we can show that a contradiction arises. Thus, $\alpha \subset \bigcup_{i,j=1}^2 A_{(i,j)}$.

We have shown that in all possible cases $\alpha \subset \bigcup_{i,j=1}^2 A_{(i,j)}$; hence, the lemma follows. \square

The following corollary follows from Lemma 5.2.

Corollary 5.3. *The set $\{(\frac{1}{6}, \frac{1}{6}), (\frac{5}{6}, \frac{1}{6}), (\frac{1}{6}, \frac{5}{6}), (\frac{5}{6}, \frac{5}{6})\}$ is a unique optimal set of four-means of the affine measure P with quantization error $V_4 = \frac{1}{36} = 0.0277778$.*

Remark 5.4. Let α be an optimal set of n -means, and $n_{ij} = \text{card}(\beta_{ij})$ where $\beta_{ij} = \alpha \cap A_{(i,j)}$ for $1 \leq i, j \leq 2$. Then, $0 \leq |n_{ij} - n_{pq}| \leq 1$ for $1 \leq i, j, p, q \leq 2$.

Lemma 5.5. *Let $n \geq 4$ and α be an optimal set of n -means for the product measure $P_c \times P_c$. For $1 \leq i, j \leq 2$, set $\beta_{ij} := \alpha \cap A_{(i,j)}$, and let $n_{ij} = \text{card}(\beta_{ij})$. Then, $U_{(i,j)}^{-1}(\beta_{ij})$ is an optimal set of n_{ij} -means, and $V_n = \sum_{i,j=1}^2 \frac{1}{36} V_{n_{ij}}$.*

Proof. For $n \geq 4$, by Lemma 5.2, we have $\alpha = \bigcup_{i,j=1}^2 \beta_{ij}$, $n = \sum_{i,j=1}^2 n_{ij}$, and so

$$V_n = \sum_{i,j=1}^2 \int_{A_{(i,j)}} \min_{a \in \beta_{ij}} \|x - a\|^2 d(P_c \times P_c).$$

If $U_{(1,1)}^{-1}(\beta_{11})$ is not an optimal set of n_{11} -means for $P_c \times P_c$, then there exists a set $\gamma_{11} \subset \mathbb{R}^2$ with $\text{card}(\gamma_{11}) = n_{11}$ such that $\int \min_{a \in \gamma_{11}} \|x - a\|^2 d(P_c \times P_c) < \int \min_{a \in U_{(1,1)}^{-1}(\beta_{11})} \|x - a\|^2 d(P_c \times P_c)$. But then, $\delta := U_{(1,1)}^{-1}(\gamma_{11}) \cup \beta_{12} \cup \beta_{21} \cup \beta_{22}$ is a set of cardinality n and it satisfies $\int \min_{a \in \delta} \|x - a\|^2 d(P_c \times P_c) < \int \min_{a \in \alpha} \|x - a\|^2 d(P_c \times P_c)$ contradicting the fact that α is an optimal set of n -means for $P_c \times P_c$. Similarly, it can be proved that $U_{(1,2)}^{-1}(\beta_{12})$, $U_{(2,1)}^{-1}(\beta_{21})$, and $U_{(2,2)}^{-1}(\beta_{22})$

are optimal sets of n_{12^-} , n_{21^-} , and n_{22^-} -means respectively. Thus,

$$\begin{aligned} V_n &= \sum_{i,j=1}^2 \frac{1}{4} \int \min_{a \in \beta_{ij}} \|x - a\|^2 d((P_c \times P_c) \circ U_{(i,j)}^{-1}) = \sum_{i,j=1}^2 \frac{1}{36} \int \min_{a \in U_{(i,j)}^{-1}(\beta_{ij})} \|x - a\|^2 dP \\ &= \sum_{i,j=1}^2 \frac{1}{36} V_{n_{ij}}, \end{aligned}$$

which gives the lemma. \square

Let us now prove the following proposition.

Proposition 5.6. *Let $n \in \mathbb{N}$ be such that $n = 4^{\ell(n)}$ for some positive integer $\ell(n)$. Then, the set*

$$\alpha_{4^{\ell(n)}} := \bigcup_{(\sigma, \tau) \in \{1, 2\}^{\ell(n)*2}} \{(A(\sigma), A(\tau))\}$$

forms a unique optimal set of n -means for the affine measure P with quantization error

$$V_{4^{\ell(n)}} = \frac{1}{4} \frac{1}{9^{\ell(n)}}.$$

Proof. We will prove the statement by induction. By Corollary 5.3, it is true if $\ell(n) = 1$. Let us assume that it is true for $n = 4^k$ for some positive integer k . We now show that it is also true if $n = 4^{k+1}$. Let β be an optimal set of 4^{k+1} -means. Set $\beta_{ij} := \beta \cap A_{(i,j)}$ for $1 \leq i, j \leq 2$. Then, by Lemma 5.2 and Lemma 5.5, $U_{(i,j)}^{-1}(\beta_{ij})$ is an optimal set of 4^k -means, and so $U_{(i,j)}^{-1}(\beta_{ij}) = \{(A(\sigma), A(\tau)) : (\sigma, \tau) \in \{1, 2\}^{k*2}\}$ which implies $\beta_{ij} = \{(A(i\sigma), A(j\tau)) : (\sigma, \tau) \in \{1, 2\}^{k*2}\}$. Thus, $\beta = \bigcup_{i,j=1}^2 \beta_{ij} = \{(A(\sigma), A(\tau)) : (\sigma, \tau) \in \{1, 2\}^{(k+1)*2}\}$ is an optimal set of 4^{k+1} -means. Since $(A(\sigma), A(\tau))$ is the centroid of $A_{(\sigma, \tau)}$ for each $(\sigma, \tau) \in I^{k+1}$, the set β is unique. Now, by Lemma 5.5, we have the quantization error as

$$V_{k+1} = \sum_{i,j=1}^2 \frac{1}{36} V_k = \frac{1}{9} \cdot \frac{1}{4} \cdot \frac{1}{9^k} = \frac{1}{4} \frac{1}{9^{k+1}}.$$

Thus, by induction, the proof of the proposition is complete. \square

Definition 5.7. *For $n \in \mathbb{N}$ with $n \geq 4$ let $\ell(n)$ be the unique natural number with $4^{\ell(n)} < n \leq 2 \cdot 4^{\ell(n)}$. For $I \subset \{1, 2\}^{\ell(n)*2}$ with $\text{card}(I) = n - 4^{\ell(n)}$ let $\alpha_n(I)$ be the set defined as follows:*

$$\alpha_n(I) = \bigcup_{(\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus I} \{(A(\sigma), A(\tau))\} \cup \left(\bigcup_{(\sigma, \tau) \in I} \{(A(\sigma 1), A(\tau)), (A(\sigma 2), A(\tau))\} \right).$$

Remark 5.8. In Definition 5.7, instead of choosing the set $\{(A(\sigma 1), A(\tau)), (A(\sigma 2), A(\tau))\}$, one can choose $\{(A(\sigma), A(\tau 1)), (A(\sigma), A(\tau 2))\}$, i.e., the set associated with each $(\sigma, \tau) \in I$ can be chosen in two different ways. Moreover, the subset I can be chosen from $\{1, 2\}^{\ell(n)*2}$ in $4^{\ell(n)} C_{n-4^{\ell(n)}}$ ways. Hence, the number of the sets $\alpha_n(I)$ is $2^{\text{card}(I)} \cdot 4^{\ell(n)} C_{n-4^{\ell(n)}}$.

The following example illustrates Definition 5.7.

Example 5.9. Let $n = 5$. Then, $\ell(n) = 1$, $I \subset \{1, 2\}^{*2}$ with $\text{card}(I) = 1$, and so

$$\begin{aligned} \alpha_5(\{(1, 1)\}) &= \{(A(1), A(2)), (A(2), A(1)), (A(2), A(2))\} \cup \{(A(11), A(1)), (A(12), A(1))\} \\ &= \left\{ \left(\frac{1}{6}, \frac{5}{6} \right), \left(\frac{5}{6}, \frac{1}{6} \right), \left(\frac{5}{6}, \frac{5}{6} \right) \right\} \cup \left\{ \left(\frac{1}{18}, \frac{1}{6} \right), \left(\frac{5}{18}, \frac{1}{6} \right) \right\}, \end{aligned}$$

or,

$$\begin{aligned} \alpha_5(\{(1, 1)\}) &= \{(A(1), A(2)), (A(2), A(1)), (A(2), A(2))\} \cup \{(A(1), A(11)), (A(1), A(12))\} \\ &= \left\{ \left(\frac{1}{6}, \frac{5}{6} \right), \left(\frac{5}{6}, \frac{1}{6} \right), \left(\frac{5}{6}, \frac{5}{6} \right) \right\} \cup \left\{ \left(\frac{1}{6}, \frac{1}{18} \right), \left(\frac{1}{6}, \frac{5}{18} \right) \right\}. \end{aligned}$$

Similarly, one can get six more sets by taking $I = \{(1, 2)\}$, $\{(2, 1)\}$, or $\{(2, 2)\}$, i.e., the number of the sets $\alpha_n(I)$ in this case is $2^{\text{card}(I)} \cdot 4^{\ell(n)} C_{n-4^{\ell(n)}} = 8$.

Proposition 5.10. *Let $n \geq 4$ and $\alpha_n(I)$ be the set as defined in Definition 5.7. Then, $\alpha_n(I)$ forms an optimal set of n -means with quantization error*

$$V_n = \frac{1}{4} \frac{1}{36^{\ell(n)}} \left(2 \cdot 4^{\ell(n)} - n + \frac{5}{9} (n - 4^{\ell(n)}) \right).$$

Proof. We have $n = 4^{\ell(n)} + k$ where $1 \leq k \leq 4^{\ell(n)}$. Set $\beta_{ij} = \alpha \cap A_{ij}$ with $n_{ij} = \text{card}(\beta_{ij})$ for $1 \leq i, j \leq 2$. Let us prove it by induction. We first assume $k = 1$. By Lemma 5.2 and Lemma 5.5, we can assume that each of $U_{(i,j)}^{-1}(\beta_{ij})$ for $i = 2$ and $j = 1, 2$, are optimal sets of $4^{\ell(n)-1}$ -means and $U_{(1,1)}^{-1}(\beta_{11})$ is an optimal set of $(4^{\ell(n)-1} + 1)$ -means. Thus, for $i = 2$ and $j = 1, 2$, we can write

$$\begin{aligned} U_{(i,j)}^{-1}(\beta_{ij}) &= \{(A(\sigma), A(\tau)) : (\sigma, \tau) \in \{1, 2\}^{(\ell(n)-1)*2}\}, \text{ and} \\ U_{(1,1)}^{-1}(\beta_{11}) &= \{(A(\sigma), A(\tau)) : (\sigma, \tau) \in \{1, 2\}^{(\ell(n)-1)*2} \setminus \{\tau\}\} \cup U_{\tau}(\alpha_2), \end{aligned}$$

for some $\tau \in \{1, 2\}^{(\ell(n)-1)*2}$, where α_2 is an optimal set of two-means. Thus

$$\alpha_n(\{(1, 1)\tau\}) = \bigcup_{i,j=1}^2 \beta_{ij} = \{(A(\sigma), A(\tau)) : (\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus \{(1, 1)\tau\}\} \cup U_{(1,1)\tau}(\alpha_2),$$

for some $\tau \in \{1, 2\}^{(\ell(n)-1)*2}$, where α_2 is an optimal set of two-means. Notice that instead of choosing $U_{(1,1)}^{-1}(\beta_{11})$ as an optimal set of $(4^{\ell(n)-1} + 1)$ -means, one can choose any one from $U_{(i,j)}^{-1}(\beta_{ij})$ for $i = 2, j = 1, 2$, as an optimal set of $(4^{\ell(n)-1} + 1)$ -means. Hence, for $n = 4^{\ell(n)} + 1$, one can write

$$\alpha_n(I) = \bigcup_{i,j=1}^2 \beta_{ij} = \{(A(\sigma), A(\tau)) : (\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus \{\tau\}\} \cup U_{\tau}(\alpha_2),$$

where $I = \{\tau\}$ for some $\tau \in \{1, 2\}^{\ell(n)*2}$ as an optimal set of n -means. Thus, we see that the proposition is true if $n = 4^{\ell(n)} + k$. Similarly, one can prove that the proposition is true for any $1 \leq k \leq 4^{\ell(n)}$. Then, the quantization error is

$$\begin{aligned} V_n &= \min_{(a,b) \in \alpha_n(I)} \|x - (a, b)\|^2 dP = \sum_{(\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus I} \int_{A_{\sigma} \times A_{\tau}} \|x - (A(\sigma), A(\tau))\|^2 d(P_c \times P_c) \\ &\quad + \sum_{(\sigma, \tau) \in I} \sum_{i=1}^2 \int_{A_{\sigma i} \times A_{\tau}} \|x - (A(\sigma i), A(\tau))\|^2 d(P_c \times P_c) \\ &= \sum_{(\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus I} \frac{1}{4^{\ell(n)}} (u_{\sigma}^2 + u_{\tau}^2) \frac{1}{8} + \sum_{(\sigma, \tau) \in I} \sum_{i=1}^2 \frac{1}{4^{\ell(n)}} \frac{1}{2} (u_{\sigma i}^2 + u_{\tau}^2) \frac{1}{8} \\ &= \sum_{(\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus I} \frac{1}{4^{\ell(n)}} (u_{\sigma}^2 + u_{\tau}^2) \frac{1}{8} + \sum_{(\sigma, \tau) \in I} \frac{1}{4^{\ell(n)}} \left(\frac{1}{9} u_{\sigma}^2 + u_{\tau}^2 \right) \frac{1}{8}. \end{aligned}$$

Since, $\text{card}(\{1, 2\}^{\ell(n)*2} \setminus I) = 2 \cdot 4^{\ell(n)} - n$, $\text{card}(I) = n - 4^{\ell(n)}$, $u_{\sigma} = u_{\tau} = \frac{1}{3^{\ell(n)}}$, upon simplification, we have $V_n = \frac{1}{4} \frac{1}{36^{\ell(n)}} \left(2 \cdot 4^{\ell(n)} - n + \frac{5}{9} (n - 4^{\ell(n)}) \right)$. Thus, the proof of the proposition is complete. \square

Definition 5.11. *For $n \in \mathbb{N}$ with $n \geq 4$ let $\ell(n)$ be the unique natural number with $2 \cdot 4^{\ell(n)} < n < 4^{\ell(n)+1}$. For $I \subset \{1, 2\}^{\ell(n)*2}$ with $\text{card}(I) = n - 2 \cdot 4^{\ell(n)}$ let $\alpha_n(I)$ be the set defined as*

follows:

$$\begin{aligned}\alpha_n(I) = & \bigcup_{(\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus I} \{(A(\sigma 1), A(\tau)), (A(\sigma 2), A(\tau))\} \\ & \cup \left(\bigcup_{(\sigma, \tau) \in I} \{(A(\sigma 1), A(\tau 1)), (A(\sigma 1), A(\tau 2)), (A(\sigma 2), A(\tau))\} \right).\end{aligned}$$

Remark 5.12. In Definition 5.11, instead of choosing the set $\{(A(\sigma 1), A(\tau)), (A(\sigma 2), A(\tau))\}$, one can choose $\{(A(\sigma), A(\tau 1)), (A(\sigma), A(\tau 2))\}$. Instead of choosing the set $\{(A(\sigma 1), A(\tau 1)), (A(\sigma 1), A(\tau 2)), (A(\sigma 2), A(\tau))\}$, one can choose either the set $\{(A(\sigma 1), A(\tau)), (A(\sigma 2), A(\tau 1)), (A(\sigma 2), A(\tau 2))\}$, or $\{(A(\sigma 1), A(\tau 1)), (A(\sigma 2), A(\tau 1)), (A(\sigma), A(\tau 2))\}$, or $\{(A(\sigma), A(\tau 1)), (A(\sigma 1), A(\tau 2)), (A(\sigma 2), A(\tau 2))\}$, i.e., the set corresponding to each $(\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus I$ can be chosen in two different ways, and the set corresponding to each $(\sigma, \tau) \in I$ can be chosen in four different ways. Since $\text{card}(\{1, 2\}^{\ell(n)*2} \setminus I) = 4^{\ell(n)} - (n - 2 \cdot 4^{\ell(n)}) = 3 \cdot 4^{\ell(n)} - n$ and the subset I can be chosen from $\{1, 2\}^{\ell(n)*2}$ in $4^{\ell(n)} C_{n-2 \cdot 4^{\ell(n)}}$ ways, the number of the sets $\alpha_n(I)$ is $2^{3 \cdot 4^{\ell(n)} - n} \cdot 4^{\text{card}(I)} \cdot 4^{\ell(n)} C_{n-2 \cdot 4^{\ell(n)}}$.

We now give an example illustrating Definition 5.11.

Example 5.13. Let $n = 9$. Then, $\ell(n) = 1$, $I \subset \{1, 2\}^{*2}$ with $\text{card}(I) = 1$. Take $I = \{(1, 1)\}$. Then,

$$\begin{aligned}\alpha_9(\{(1, 1)\}) &= \{(A(11), A(2)), (A(12), A(2)), (A(21), A(2)), (A(22), A(2)), (A(21), A(1)), \\ &\quad (A(22), A(1))\} \cup \{(A(11), A(1)), (A(12), A(1)), (A(12), A(12))\} \\ &= \left\{ \left(\frac{1}{18}, \frac{5}{6} \right), \left(\frac{5}{18}, \frac{5}{6} \right), \left(\frac{13}{18}, \frac{5}{6} \right), \left(\frac{17}{18}, \frac{5}{6} \right), \left(\frac{13}{18}, \frac{1}{6} \right), \left(\frac{17}{18}, \frac{1}{6} \right) \right\} \\ &\quad \cup \left\{ \left(\frac{1}{18}, \frac{1}{6} \right), \left(\frac{5}{18}, \frac{1}{6} \right), \left(\frac{5}{18}, \frac{5}{18} \right) \right\}.\end{aligned}$$

Note that each of $\alpha_9(\{(1, 1)\})$, $\alpha_9(\{(1, 2)\})$, $\alpha_9(\{(2, 1)\})$, $\alpha_9(\{(2, 2)\})$ can be chosen in 32 ways, i.e., the numbers of the sets $\alpha_9(I)$ in this case is $4 \cdot 32 = 128$. Moreover, using the formula in Remark 5.12, we have

$$2^{3 \cdot 4^{\ell(n)} - n} \cdot 4^{\text{card}(I)} \cdot 4^{\ell(n)} C_{n-2 \cdot 4^{\ell(n)}} = 128.$$

Proposition 5.14. Let $n \geq 4$ and $\alpha_n(I)$ be the set as defined in Definition 5.11. Then, $\alpha_n(I)$ forms an optimal set of n -means with quantization error

$$V_n = \frac{1}{36^{\ell(n)+1}} (9 \cdot 4^{\ell(n)} - 2n).$$

Proof. We have $n = 2 \cdot 4^{\ell(n)} + k$ where $1 \leq k < 2 \cdot 4^{\ell(n)}$. Set $\beta_{ij} = \alpha \cap A_{ij}$ with $n_{ij} = \text{card}(\beta_{ij})$ for $1 \leq i, j \leq 2$. Let us prove it by induction. We first assume $k = 1$. By Lemma 5.2 and Lemma 5.5, we can assume that each of $U_{(i,j)}^{-1}(\beta_{ij})$ for $i = 2$ and $j = 1, 2$, are optimal sets of $2 \cdot 4^{\ell(n)-1}$ -means and $U_{(1,1)}^{-1}(\beta_{11})$ is an optimal set of $(2 \cdot 4^{\ell(n)-1} + 1)$ -means. Thus, for $i = 2$ and $j = 1, 2$, we can write

$$\begin{aligned}U_{(i,j)}^{-1}(\beta_{ij}) &= \{U_{(\sigma, \tau)}(\alpha_2) : (\sigma, \tau) \in \{1, 2\}^{(\ell(n)-1)*2}\}, \text{ and} \\ U_{(1,1)}^{-1}(\beta_{11}) &= \{U_{(\sigma, \tau)}(\alpha_2) : (\sigma, \tau) \in \{1, 2\}^{(\ell(n)-1)*2} \setminus \{\tau\}\} \cup U_{\tau}(\alpha_3),\end{aligned}$$

for some $\tau \in \{1, 2\}^{(\ell(n)-1)*2}$, where α_3 is an optimal set of three-means. Thus

$$\alpha_n(\{(1, 1)\tau\}) = \bigcup_{i,j=1}^2 \beta_{ij} = \{U_{(\sigma, \tau)}(\alpha_2) : (\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus \{(1, 1)\tau\}\} \cup U_{(1,1)\tau}(\alpha_3),$$

for some $\tau \in \{1, 2\}^{(\ell(n)-1)*2}$, where α_3 is an optimal set of three-means. Notice that instead of choosing $U_{(1,1)}^{-1}(\beta_{11})$ as an optimal set of $(2 \cdot 4^{\ell(n)-1} + 1)$ -means, one can choose any one

from $U_{(i,j)}^{-1}(\beta_{ij})$ for $i = 2, j = 1, 2$, as an optimal set of $(2 \cdot 4^{\ell(n)-1} + 1)$ -means. Hence, for $n = 2 \cdot 4^{\ell(n)} + 1$, one can write

$$\alpha_n(I) = \bigcup_{i,j=1}^2 \beta_{ij} = \{U_{(\sigma,\tau)}(\alpha_2) : (\sigma, \tau) \in \{1, 2\}^{\ell(n)*2} \setminus \{\tau\}\} \cup U_\tau(\alpha_3),$$

where $I = \{\tau\}$ for some $\tau \in \{1, 2\}^{\ell(n)*2}$ as an optimal set of n -means. Thus, we see that the proposition is true if $n = 2 \cdot 4^{\ell(n)} + 1$. Similarly, one can prove that the proposition is true for any $1 \leq k < 2 \cdot 4^{\ell(n)}$. Thus, writing $\alpha_2 = \{(A(1), A(\emptyset)), (A(2), A(\emptyset))\}$, and $\alpha_3 = \{(A(1), A(1)), (A(1), A(2)), (A(2), A(\emptyset))\}$, we have, in general,

$$\begin{aligned} \alpha_n(I) = & \bigcup_{(\sigma,\tau) \in \{1,2\}^{\ell(n)*2} \setminus I} \{(A(\sigma 1), A(\tau)), (A(\sigma 2), A(\tau))\} \\ & \cup \left(\bigcup_{(\sigma,\tau) \in I} \{(A(\sigma 1), A(\tau 1)), (A(\sigma 1), A(\tau 2)), (A(\sigma 2), A(\tau))\} \right), \end{aligned}$$

where $I \subset \{1, 2\}^{\ell(n)*2}$ with $\text{card}(I) = k$ for some $1 \leq k < 2 \cdot 4^{\ell(n)}$. Then, we obtain the quantization error as

$$\begin{aligned} V_n = & \min_{(a,b) \in \beta_n(I)} \int \|x - (a, b)\|^2 dP = \sum_{(\sigma,\tau) \in \{1,2\}^{\ell(n)*2} \setminus I} \sum_{i=1}^2 \int_{A_{\sigma i} \times A_\tau} \|x - (A(\sigma i), A(\tau))\|^2 d(P_c \times P_c) \\ & + \sum_{(\sigma,\tau) \in I} \left(\sum_{j=1}^2 \int_{A_{\sigma 1} \times A_{\tau j}} \|x - (A(\sigma 1), A(\tau j))\|^2 d(P_c \times P_c) \right. \\ & \left. + \int_{A_{\sigma 2} \times A_\tau} \|x - (A(\sigma 2), A(\tau))\|^2 d(P_c \times P_c) \right) \\ = & \sum_{(\sigma,\tau) \in \{1,2\}^{\ell(n)*2} \setminus I} \sum_{i=1}^2 \frac{1}{4^{\ell(n)}} \frac{1}{2} (u_{\sigma i}^2 + u_\tau^2) \frac{1}{8} + \sum_{(\sigma,\tau) \in I} \frac{1}{4^{\ell(n)}} \left(\sum_{j=1}^2 \frac{1}{4} (u_{\sigma 1}^2 + u_{\tau j}^2) \frac{1}{8} + \frac{1}{2} (u_{\sigma 2}^2 + u_\tau^2) \frac{1}{8} \right) \\ = & \sum_{(\sigma,\tau) \in \{1,2\}^{\ell(n)*2} \setminus I} \frac{1}{4^{\ell(n)}} \left(\frac{1}{9} u_\sigma^2 + u_\tau^2 \right) \frac{1}{8} + \sum_{(\sigma,\tau) \in I} \frac{1}{4^{\ell(n)}} (u_\sigma^2 + 5u_\tau^2) \frac{1}{72}. \end{aligned}$$

Since, $\text{card}(\{1, 2\}^{\ell(n)*2} \setminus I) = 3 \cdot 4^{\ell(n)} - n$, $\text{card}(I) = n - 2 \cdot 4^{\ell(n)}$, $u_\sigma = u_\tau = \frac{1}{3 \cdot 4^{\ell(n)}}$, upon simplification, we have $V_n = \frac{1}{36^{\ell(n)+1}} (9 \cdot 4^{\ell(n)} - 2n)$. Thus, the proof of the proposition is complete. \square

Remark 5.15. It is well-known that the optimal set of one-mean for a probability distribution is always the expected value of the distribution and the corresponding quantization error is the variance. Propositions 3.3, 3.4, 5.6, 5.10 and 5.14 give all the optimal sets of n -means for all $n \geq 2$, and the corresponding quantization error for the affine measure P .

REFERENCES

- [1] E.F. Abaya and G.L. Wise, *Some remarks on the existence of optimal quantizers*, Statistics & Probability Letters, Volume 2, Issue 6, December 1984, pp. 349-351.
- [2] D. Çömez and M.K. Roychowdhury, *Optimal quantizers for probability distributions on Sierpinski carpets*, arXiv:1511.01990 [math.DS].
- [3] Q. Du, V. Faber and M. Gunzburger, *Centroidal Voronoi Tessellations: Applications and Algorithms*, SIAM Review, Vol. 41, No. 4 (1999), 637-676.
- [4] A. Gersho and R.M. Gray, *Vector quantization and signal compression*, Kluwer Academy publishers: Boston, 1992.
- [5] R.M. Gray, J.C. Kieffer and Y. Linde, *Locally optimal block quantizer design*, Information and Control, 45 (1980), 178-198.

- [6] A. György and T. Linder, *On the structure of optimal entropy-constrained scalar quantizers*, IEEE transactions on information theory, vol. 48, no. 2, February 2002.
- [7] S. Graf and H. Luschgy, *Foundations of quantization for probability distributions*, Lecture Notes in Mathematics 1730, Springer, Berlin, 2000.
- [8] S. Graf and H. Luschgy, *The Quantization of the Cantor Distribution*, Math. Nachr., 183 (1997), 113-133.
- [9] R. Gray and D. Neuhoff, *Quantization*, IEEE Trans. Inform. Theory, 44 (1998), 2325-2383.
- [10] J. Hutchinson, *Fractals and self-similarity*, Indiana Univ. J., 30 (1981), 713-747.
- [11] M. Moran, *Hausdorff measure of infinitely generated self-similar sets*, Monatsh. Math. 122, 1996, 387-399.
- [12] D. Mauldin and M. Urbański, *Dimensions and measures in infinite iterated function systems*, Proc. London Math. Soc. (3) 73 (1996), no. 1, 105-154.
- [13] L. Roychowdhury, *Optimal quantizers for probability distributions on nonhomogeneous Cantor sets*, arXiv:1512.00379 [stat.CO].
- [14] M.K. Roychowdhury, *Optimal quantization for infinite similitudes on the real line*, arXiv:1512.09161 [math.DS].
- [15] L. Roychowdhury and M.K. Roychowdhury, *Optimal quantization for infinite nonhomogeneous distributions on the real line*, arXiv:1603.00731 [math.DS].
- [16] M.K. Roychowdhury, *Quantization and centroidal Voronoi tessellations for probability measures on dyadic Cantor sets*, arXiv:1509.06037 [math.DS].

DEPARTMENT OF MATHEMATICS, 408E24 MINARD HALL, NORTH DAKOTA STATE UNIVERSITY, FARGO, ND 58108-6050, USA.

E-mail address: `Dogan.Comez@ndsu.edu`

SCHOOL OF MATHEMATICAL AND STATISTICAL SCIENCES, UNIVERSITY OF TEXAS RIO GRANDE VALLEY, 1201 WEST UNIVERSITY DRIVE, EDINBURG, TX 78539-2999, USA.

E-mail address: `mrinal.roychowdhury@utrgv.edu`